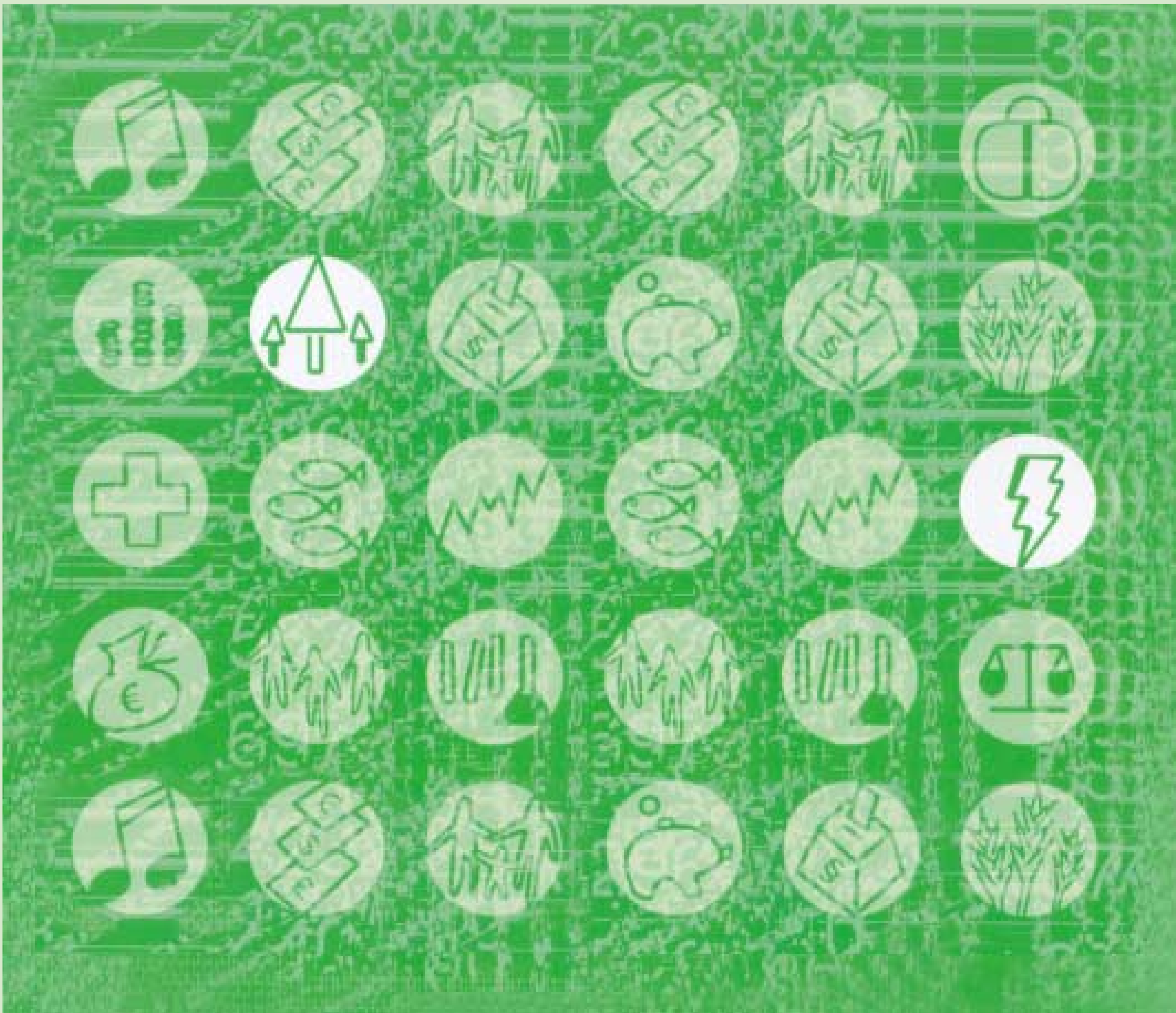
Tilastokeskus
Statistikcentralen
Statistics Finland

Reviews 2006/2

Timo Koskimäki – Mari Ylä-Jarkko – Mari Kinnunen (Eds.)

International Working Group on Price Indices – The Ottawa Group

Proceedings of the Eighth Meeting Helsinki, August 2004



Tilastokeskus
Statistikcentralen
Statistics Finland

Reviews 2006/2

Timo Koskimäki – Mari Ylä-Jarkko – Mari Kinnunen (Eds)

International Working Group on Price Indices – The Ottawa Group

Proceedings of the Eighth Meeting Helsinki, August 2004

Helsinki 2006

Inquiries:

Timo Koskimäki
(09) 17 341

Cover photograph: Rurik Mahlberg

ISSN 1239–3800
ISBN 952–467–530–7

Helsinki 2006

Preface

In August 2004 Statistics Finland hosted the 8th meeting of the Ottawa Group, the international working group on price indices. This publication contains the post-conference final versions of the papers that were presented in the meeting as well as the session summaries prepared by the session moderators. In addition to this publication, these papers will in the future also be available through the Ottawa Group web-site (www.ottawagroup.org).

On behalf of Statistics Finland I would like to express my gratitude and thanks to the session leaders, speakers and participants of the meeting for the excellent papers and for the fruitful discussion during the meeting.

Kari Molnar
Director, Prices and wages statistics

Contents

Opening address[1]

Hosting this conference is a great pleasure for Statistics Finland. Statistics Finland has, throughout the years, attempted to produce high quality official statistics that is based on up-to-date international standards and academic research. This applies particularly to the field of price indices, where we have contributed to international discussion by or own research work as well. One important background for our activity in the field of price index research is the long price index research tradition at the University of Helsinki, initiated by professor Leo Törnqvist in 1930's and carried forward by his students and successors. You will be hearing some recent results of this tradition also during this meeting. I hope you will enjoy these and all the other interesting papers to be presented during the next days.

Let me now say a few words about the role of the Ottawa city Group in general and the themes of this meeting in particular.

I think the Ottawa Group has to a large extent justified its existence as a fruitful forum for the exchange of experiences in academic research and practical CPI compilation. The most important feature, from the point of view of a statistical agency, is the city groups' capacity to propose recommendations for best practices and identify the future fields of research. It is evident, that the discussions and recommendations of the Ottawa Group have greatly influenced the contents of the new international CPI and PPI manuals. As such, I consider the publication of these two manuals as a major improvement in the field of price statistics. I would very much like to see the Ottawa Group stress the function of making recommendations in its activities, especially now as the new international manuals are both meant to be "living documents". The Ottawa group would be a natural, already existing forum to take care of the identification of the updates needed to the manuals, especially the CPI manual.

The price indices, especially CPI's, are always under a strict public watch. This is a natural consequence of the large-scale economic impact the indices have to the public and private economy. During the past decade or so, much of the price index research have been inspired by the arguments put forward in the Boskin report of 1996. The focus of research has generally been in the field of quality change issues like technical change of the technology-intensive products and, on the other hand, in the field of different index formulae designed to measure "the change in the cost of living" or "pure price change".

If we take a look at the recent public discussion relating to CPI's, the focus of the discussion seem to be switching to other areas. In Europe, in connection with the introduction of the common currency Euro, a severe distrust towards the results of the official consumer price indices have emerged. To take just one example: The first CPI results published by Statistics Finland after the introduction of the Euro pointed out that the introduction of the new currency had no effect to the general price level. After the release of the CPI results, a newspaper carried out a readers' poll posing the question: "Statistics Finland claims, that the introduction of the Euro has not caused a rise in price level. Do you agree or disagree?" Only 6 per cent of the respondents agreed, 94 per cent said that they disagree with the official CPI. Although the discussion in Finland seems to have calmed down, in several European countries the public credibility of the CPI's still suffers from the distrust that emerged in connection with the Euro changeover.
Around the world, there has also been public concerns relating to the general quality of the CPI's that have been caused by statistical treatment of some difficult areas like insurance services or

[1] Jarmo Hyrkkö, Associate Director General, Statistics Finland

housing. The introduction of new data collection methods like the use of scanner data have in some cases lead to errors in production process of the CPI's and, in the extreme cases, called for a revision of the CPI results.

The past research on Consumer Price Indices has been concentrated on strictly CPI-related issues like quality change from consumers' point of view or cost-of-living related discussion motivated by the use of CPI's as income escalators. However, from a methodological point of view, CPI's share a lot with other price indices like PPI's or international price level comparisons. There is clearly a possibility of utilising CPI-related research in the other fields of price indices as well or, more generally, to conduct research and development work that would be applicable in several price indices. The possibility of taking into account not only the CPI but also price indices in general in the work of the Ottawa Group is already inherent in the remit of the group.

I am glad to note that the agenda of the meeting does respond to the public concerns about the quality of CPI's and other developments in the field of price indices that I have pointed out. I'm looking forward to see the results of this meeting with respect to the topics in the agenda.

I would like to thank all of you that have produced a research paper for this meeting for your effort, the steering committee of the Ottawa Group that has supervised the organisation of this meeting, those of you that have volunteered to chair sessions and produce summaries of the recommendation and, of course, all of you that have devoted your time and come here to discuss price index issues.

Session 1 - Price indices for services

Moderator: Rósmundur Guðnason, Statistics Iceland

Summary of Session

In this session three papers were presented:

George Beelen described the treatment of property-casualty insurance in the Canadian CPI, the three categories of such insurance being vehicle, homeowners' and tenants' insurance. The current treatment of insurance was explained in terms of scope, weight and pricing with comparisons to the discussion of insurance in the new international CPI manual.

There has been significant public attention on property-casualty insurance in the Canadian CPI. All insurance categories have increased considerably more than the All-items CPI in the last few years.

The Canadian CPI was criticized as using inappropriate concepts and inconsistent methodology for property-casualty insurance. Some of these criticisms were that the Canadian CPI basket placed too much weight on insurance, that the price change should be not be based on changes in gross premiums concept but rather on changes in "premiums less claims', and that the CPI should hold the probability of occurrence of risk events constant in its pricing methods.

The paper argues that even if a direct method of holding all risk factors constant could be developed, it would not necessarily be desirable for the purposes of the official Canadian CPI, based on a risk-assuming definition of insurance services. Tracking changes in the gross premiums of insurance policies of constant terms and conditions without adjusting for risk probabilities is the most appropriate method for a CPI whose primary purpose is to reflect changes in the purchasing power of the consumer dollar.

David Fenwick presented a paper by Fenwick and Ball dealing with different strategies for the pricing of mobile telecom services. The traditional tariff approach to the costing of mobile telephone calls fails to take into account the full range of package options available to consumers over the life of the "fixed" basket. The method proposed in this paper samples profiles and then selects for pricing the appropriate package from each service provider. It assumes a rational consumer who has perfect knowledge but this assumption is more realistic than the alternative of the ill-informed laggard, particularly in telecommunications in the UK where unit cost information is readily available and where at points in time customers can transfer from one tariff to another free of charge."

Keith Woolford presented an experimental index ABS aims to use fro measuring price changes for some of the most significant financial services acquired by households — deposit and loan facilities provided by financial institutions and services associated with the acquisition and disposal of shares and real estate.

Recommendation for Statistical Agencies:

The production of reliable price measures for services like insurance or financial services is not straightforward. The development of proper measurement systems is costly and often requires intensive co-operation with companies and institutions providing such services. Statistical agencies are encouraged to foresee this when planning to undertake development and research work in this important area.

George Beelen[2]: The Treatment of Property-Casualty Insurance in the Canadian CPI[3]

Introduction and Background: This paper is about the treatment of property-casualty insurance in the Canadian CPI, the three categories of such insurance being vehicle, homeowners' and tenants' insurance. The current treatment of insurance is explained in terms of scope, weight and pricing with comparisons to the discussion of insurance in the new international CPI manual.

The paper was motivated mainly by two events. First, Statistics Canada is reviewing its methods in light of the new CPI manual. Second, there has been significant public attention on property-casualty insurance in the Canadian CPI.

All insurance categories have increased considerably more than the All-items CPI in the last few years. Insurance industry sources have attributed rising vehicle premium rates to increased claims costs, including higher pain and suffering awards made in the justice system, higher hospital care costs (described as much higher "human body repair" than "auto body repair" costs) , and, increased rates of fraudulent claims. Lower returns on investment of premiums (i.e. premium supplements) were also cited as a factor driving up premium rates by some analysts and commentators, although industry representatives have stated that this is not a major factor.

The criticisms of the Canadian CPI arose in part because timing problems caused some premium rate increases to be reflected in the CPI later than when the rate increases actually occurred, particularly for vehicle insurance. Lags in the provision of information to Statistics Canada on premium rate changes by some companies occurred. As a result, the premium rate increases that occurred in fact over a three year period were compressed into about an 18-month time period in the CPI. These timing problems should be avoided in future based on recent work with the insurance industry association and with individual insurance company respondents to ensure better and timelier reporting.

The Canadian CPI was also criticized as using inappropriate concepts and inconsistent methodology for property-casualty insurance. Some of these criticisms were that the Canadian CPI basket placed too much weight on insurance, that the price change should be not be based on changes in gross premiums concept but rather on changes in "premiums less claims', and that the CPI should hold the probability of occurrence of risk events constant in its pricing methods.

Scope in the Canadian CPI: Property-casualty insurance (homeowners', tenants', vehicle) categories are currently included in the Canadian CPI. Homeowners' insurance is included in the official Canadian CPI since a homeowner-cost concept rather than rental equivalence method is used for owner-occupied housing.[4]

Whole life insurance is excluded from the CPI, mainly because the investment element of whole life insurance cannot be separated from the term insurance component. Term life and disability insurance are currently not included in the Canadian CPI due to the practical difficulties of producing suitable price indexes rather than for any conceptual reasons. In practice some disability insurance

[2] Statistics Canada
[3] This paper reflects the views of the author and not necessarily those of Statistics Canada
[4] However, for analytical and comparison purposes, data are published regularly using other conceptual treatments for owner-occupied housing including rental equivalence, money outlays and net purchase concepts.

is in fact reflected in the CPI via the vehicle insurance component since there is usually a disability insurance component in vehicle insurance premiums.

Health services received by the population via the publicly-funded health "insurance" system are excluded from the CPI. In some provinces these services are completely financed by taxes and therefore are treated as a public service. In other provinces, households pay public health "insurance premiums" but these are considered similar to income taxes (and in fact are often administered similarly to income taxes) and thus are outside the scope of the Canadian CPI.

Private supplementary health, pharmaceutical and dental insurances are not included in the CPI but are not excluded on conceptual grounds. To date, they are not reflected in the CPI due to the practical difficulties in producing price indexes for these services.

Scope in the new CPI manual: The concepts and scope chapter in the new CPI manual discusses insurance in paragraph 3.49. "In the case of non-life insurance, the net premium is essentially a transfer that goes into a pool covering the collective risks of policy holders as a whole. As a transfer, it falls outside the scope of a CPI. In the case of life insurance, the net premium is essentially a form of financial investment. It constitutes the purchase of a financial asset, which is also outside the scope of a CPI".

The manual therefore considers the portion of premiums used to fund claims to be a transfer and outside the scope of the CPI. This would imply that the weights for non-life insurance in a CPI should exclude the portion of gross premiums that funds claims. This treatment apparently derives from a risk-pooling view of insurance, i.e. provision of insurance administration services on behalf of a population. As will be explained in this note, this is not consistent with the current treatment of non-life insurance in the Canadian CPI.

It seems that the subsequent discussion of insurance in the CPI manual in Chapter 10 is also not consistent with the scope statement for insurance in Chapter 3. If the portion of premiums used to fund claims is out of scope as stated in Chapter 3, then the only conceptually consistent weighting basis for insurance is premiums net of claims.

Weight of Property-Casualty Insurance in the Canadian CPI: The CPI for Canada uses gross expenditures on property-casualty insurance premiums of households as reported to the Survey of Household Spending or SHS (up to 1996 called the Family Expenditure Survey) as the basis of weights for vehicle, tenants' and homeowners' insurance categories. In the most recent basket update based on expenditures in 2001, vehicle insurance has about 2.7% of the total basket weight, homeowners' (home and contents) insurance about 1% and tenants' insurance 0.1%.

Insurance is weighted on the basis of gross premiums and expenditures on all other goods and services are supposed to be net of consumers' purchases financed by insurance claims (this is the option in the CPI manual, paragraphs 10.159 and 10.160). In the Canadian SHS, households are asked to include expenditures on goods or services, and any "deductible" or co-insurance amounts paid by consumers but excluding the portion of purchases paid for or financed from claims. This approach is used to avoid reporting problems when the amounts financed by claims are not known to the consumer. For example, in Canada it is not unusual that the costs of vehicle repairs are sometimes negotiated and paid directly by an insurance company to the vehicle repair garage. However, some respondents to the SHS cannot or do not always remember to exclude claims-financed expenditures so tradeoffs exist no matter how the questions are asked.

Note that expenditures on some forms of insurance associated with other specific commodities such as travel insurance or the purchase of extended warranties (a form of insurance) on vehicles or appliances are currently included in the CPI with the expenditure weights for these products. In practice, the product specifications for pricing purposes usually do not include insurance or extended warranties, at least for products for which only a minority of consumers buy such insurance or extended warranties. In effect, changes in the prices of the commodities themselves are used as proxies for changes in the prices of the associated insurance services.

Weights in the CPI Manual: The CPI manual discusses options for weighting insurance under the payments, use and acquisitions approaches in Chapter 10. Three plausible options are: 1. Gross premiums, net expenditures 2. Net premiums, gross expenditures and 3. Gross premiums, gross expenditures. The manual points out that not all proceeds from claims are used to replace or repair damaged items. Households can use the claims funds for other purposes and some claims are paid to non-households (businesses or government) to compensate for damage or destruction to their property, for example.

The Canadian CPI uses the first option (gross premiums, net expenditures) despite these problems. The property insurance weights in the Canadian CPI reflect gross premiums paid but weights for all other goods and services reflect expenditures reduced to the extent that Canadian households consider that insurance claims proceeds were used to purchase other goods and services. To date the third option (gross premiums, gross expenditures) has not been adopted because this option has been viewed as "double-counting" expenditures since it includes both gross insurance premiums as well as expenditures on other goods and services financed in part by the gross premiums. However, the gross premiums, gross expenditures option has some advantages and will be considered for future basket updates. The main advantage of this option is that response to the Survey of Household Spending would be simpler, avoiding problems of inconsistent exclusions of claims-funded expenditures.

Measurement of Price Change for Insurance in the Canadian CPI: In the case of property insurance, the Canadian CPI measures price change by tracking gross premium rates for a sample of insurance policies. For example, for vehicle insurance, the method uses samples of insurance companies, of geographic regions, of driver profiles (age, gender and usage of vehicle and driving records), and of classes of vehicles, for insurance policies with fixed characteristics (such as deductible levels, any other co-insurance provisions, liability maximums, and other policy provisions). The policy characteristics and samples are updated periodically as judged necessary. Statistics Canada has recently updated these samples, based partly on data and advice from the Insurance Bureau of Canada.

No attempt is made to hold constant total expected claims costs of insurers. Adjustments are not made for changes in the probabilities of claims nor for "changes in the prices and repair costs of insured goods and altered health service prices" (Ralph Turvey, see Appendix of this paper). Therefore, the "quality" which is being held constant is the "security of the consumer, rather than the volume of the repairs, replacement and health treatment paid for by claims." (Turvey)

In the case of homeowners' insurance, the index measures changes in the cost of insuring a fixed stock of dwellings against a specified list of risk events or perils. This cost varies not only with changes in insurance rates for given property values, but also with changes in the values of properties covered which result from changes in dwelling prices. Thus, the price of homeowners' insurance is estimated from the product of two component index series. The first measures the change in the value of properties as measured by the house component of the New Housing Price Index. The second component index measures the change in insurance rates of identical insurance policies for properties of given constant value, using premium rate information from insurance companies in the sample.

Discussion of the Current Approach: The insurance price indexes in the Canadian CPI reflect changes in the price to consumers of obtaining insurance coverage against the economic consequences of a specified list of perils, based on changes in gross premiums for a sample of identical (or at least very similar) insurance policies. By comparing changes in premiums of identical insurance policies, the methods used ensure that the Canadian CPI does not reflect as price change the impacts of any changes in the type of property or in the population being insured. For example, the vehicle insurance price index will not change due to changes in the mix of vehicles or demographic profile of the driving population.

Insurance premiums provide the funds from which claims are paid out. Premiums also compensate insurance companies for the services they provide (the 'implicit service charge'). These insurance services include setting premium rates, collecting premiums, investing premiums (i.e. earning the "premium supplements"), and settling claims fairly (including minimizing fraudulent claims).

The current Canadian method reflects any change in gross premiums for an identical policy (i.e. with all contract details and conditions including driver and vehicle classes held constant) as a price change, whatever the reason for the change. Premiums for policies can change as a result of changes in expected claim values (arising from changes in the probabilities of the occurrence of risk events or in the value of claims arising from risk events occurring) or changes in the implicit service charges of insurance companies for administering insurance.

The method used is partly based on pragmatism. The gross premium rates for insurance policies are observable and, despite the recent timing problems, usually available in the time frame required for a CPI. However there is no direct method of holding constant some or all of the expected claims costs. In other words, there is no method available for determining whether a change in premiums is due to changes in the level of expected claims, changes in "premium supplements" or changes in the implicit service charge.

Consider the range of factors affecting risks which can cause changes in premiums in vehicle insurance for example. Expected claims costs are affected by changes in the probabilities of the perils and in the likelihood that a claim will be made, and in the level of payouts for claims. Changes in the probabilities of the occurrence of risk events are the result of many factors including highway and vehicle design, the amount of traffic congestion, even levels of enforcement of speeding laws. The probability of a claim being made is affected by changes in the willingness of potential claimants to make, or not make, a claim when legitimately entitled to do so. Drivers often face the choice of making a claim and risking personal rate increases or not making a claim and keeping a clean record with their insurer. As well, the probability of claims is affected by the rate of fraudulent claims. The average payout level of claims is affected by many factors, including changes in replacement and repair costs of property insured, health care costs, income levels (affecting payout for time lost from work) and even changes in the justice system in making awards for pain and suffering claims.

Jack Triplett[5] calls the two views of insurance the risk-assuming and risk-pooling views (the latter reflected in the 1993 System of National Accounts). The scope statement discussed earlier takes the risk-pooling view. "In the risk-assuming view of insurance, the policy holder buys the service of having assets or income protected against loss" (Triplett). The Canadian CPI treatment is consistent with the risk-assuming view of insurance.

It is how insurance services are defined that becomes the key issue in the debate in its CPI treatment. In effect the Canadian treatment of property insurance views insurance as a form of "re-

[5] "*Chapter Six: Price, Output and Productivity of Insurance: Conceptual Issues in* Services" in Productivity in the United States, New Sources of Economic Growth, Jack Triplett, Brookings Institution, pending.

placement guarantee", that is, property insurance from the consumer point of view is "considered the price of a guarantee that goods will be restored or replaced up to specified limits, in cases such as accident, fire ant theft."[6] Risks transferred to insurers include the economic costs arising from loss or damage to property, the costs of health care and rehabilitation, loss of income due to personal injury arising from an accident, and liability for property damage or injuries to others. Change in price of insurance becomes any change in the price of providing the replacement guarantee within the terms and conditions of the insurance policy.

Much debate at Statistics Canada has been on the question of what should be "held constant" for insurance. The current treatment in the CPI is said to hold constant the "level of security".

Bohdan Schultz[7] suggested approaching the problem from the perspective of the services rendered to the consumer; for example vehicle insurance as a component of private transportation services. The Canadian CPI has ten component indexes that comprise an index for private transportation services, including purchase and leasing of automobiles, gasoline, vehicle parts and supplies, maintenance and repair services, and insurance. From this perspective, any change in vehicle insurance premiums for an identical policy, whether due to changes in the probabilities of claims, in the average costs of claims or in the implicit service charges of insurance companies, would not change the amount or quality of private transportation service received by consumers and thus should be considered price change.

The opposing risk-pooling view is that any changes in premiums due to changes in expected levels of claims should not be considered to be price change but rather a change in the size of the pool of funds needed to pay claims. If insurance is the administering of the pool of funds, then changes in premiums to provide the pool of funds are not price changes.

The Purposes of the CPI and the Treatment of Insurance: In this situation of two quite different views of the definition of insurance services, we consider the question in light of the purposes of the CPI. While the official Canadian CPI is a multi-purpose indicator, the primary role of the Canadian CPI is stated to be "to provide an adequate indicator of price-induced changes in the purchasing power of the consumer dollar".[8] This has been paraphrased as "to estimate price-induced changes in living costs, a role which is frequently referred to as the escalator function.[9] An index using gross premiums for weights and price change rather than a 'net of claims' measure seems more consistent with this primary purpose of the Canadian CPI as an escalation index.

If the primary purpose were different, other methods for weights would seem to be more appropriate. Keith Woolford of ABS has stated that "changes in gross premiums alone would be relevant in the construction of an index designed to measure changes in household living costs because the amounts received by way of claims are not relevant as they are regarded as sources of income. Another way of looking at this is …in constructing an inflation index, it is important that each item in the index is accorded a weight which reflects its real economic significance, while in constructing a living cost index it is important that each item be accorded a weight which reflects its relative share of gross household expenditures".[10] The Australian Bureau of Statistics defines the principal purpose of the Australian CPI as to provide a general measure of price inflation for the household sector as a whole. The Australian CPI uses a net 'premiums less claims' weight for insurance but movements in gross insurance premiums as the price measure.

[6] Canadian CPI reference manual, 1957 and subsequent editions.
[7] "Alternative Estimates of Price Change for Private Transportation", Bohdan Schultz, 5th Meeting of Ottawa Group, 1999, Reykjavik Iceland. The point is also mentioned in Triplett, op. cit. Section V.
[8] CPI Reference paper
[9] Bohdan Schultz, "Treatment of Owner-Occupied Accommodation in the CPI", Discussion Paper, 2001.
[10] "Treatment of Insurance Services in the Australian CPI, ABS Website.

So it is argued that a gross premiums weight is appropriate for an escalation index. But what does that imply that pricing insurance for an escalation index? Adherents of the risk-assuming view argue that changes in the purchasing power of the consumer to buy a fixed level of security are best measured by changes in gross premiums.

But supporters of the risk-pooling definition of insurance in Statistics Canada argue that changes in premiums arising from changes in the size of the risk pool are not related to the price of insurance services at all. Even if the weight is based on gross premiums, the risk-pooling definition of insurance implies a pricing measure based on the implicit service charge only.

Discussion of the New CPI Manual on the Pricing of Insurance: The manual in Chapter 10 discusses pricing gross insurance premiums in paragraphs 10.167-170. The CPI manual appears to take the position that it would be desirable to make quality adjustments for changes in risk levels but in practice this is usually not done due to the practical difficulties.

10.168 While it is clear that pricing to constant quality requires these conditions to be held fixed, there is also a question about whether the risk of a claim being made should be held constant. In other words, if the incidence of, say, vehicle theft increases, should this be regarded as a quality improvement or simply a price change? If, on the one hand, it is argued that as the consumers' decision to insure is based on their assessment of the likelihood of suffering a loss compared to the premium charged, the risk factors should be held constant. On the other hand, it may be argued that, once insured, the consumer simply expects to be compensated for any loss. From the perspective of the consumer, any increase in risk simply represents an increase in the insurer's cost base (which may or may not be passed on to the consumer by way of a price change). Obtaining data of sufficient reliability to make quality adjustments in response to changes in risk is problematic, so in practice most indices reflect changes in risk as a price change.

This paper suggests that it is not necessarily desirable to make quality adjustments for changes in risk, depending on the definition of insurance service and the purpose of the index.

In any case, it appears that no practical price measurement methodology for the implicit service charge component of insurance has yet been developed for use in a CPI. There is simply no directly observable price for the implicit service charge element of insurance. In fact, if the implicit service charge were the target concept for measurement, one could argue that an administration or management services index would be a reasonable pricing basis to use.[11]

Gross premiums of a consistent set of insurance policies are directly and objectively observable and so far this approach seems to be the only practical option for CPI measurement purposes. Even in countries where the implicit service charge is the target concept for insurance in the CPI, gross premium rates are used as a proxy price indicator. There are significant advantages in public acceptance and credibility for the CPI to use pricing methods based on observable transactions. The "implicit service charge" approach to insurance in the CPI would suffer from a lack of identification by consumers with the definition of the product being priced in the basket.

A Short Digression on Fraudulent Claims and the Implicit Service Charge

The implicit service charge component of insurance is the "fee charged for calculating the risks, determining the premiums, administering the collection and investment of premiums, and the pay-

[11] Z. Griliches made this point. See Section V of Triplett, "*Chapter Six: Price, Output and Productivity of Insurance: Conceptual Issues in* Services" in Productivity in the United States, New Sources of Economic Growth, Brookings Institution, pending.

ment of claims (paragraph 3.47 in CPI manual). Insurance services are calculated as gross premiums plus premium supplements less claims less changes in actuarial reserves.

Should the implicit service charge concept exclude fraudulent claims? Honest consumers would want (and expect) insurers to minimize the paying of fraudulent claims as part of the service of administering insurance. If the rate of fraudulent claims were to increase causing an increase in premium rates, any price measure of the implicit insurance service to consumers should increase as well. However, the definition of the implicit service charge seems to exclude all claims, legitimate and fraudulent alike. I would argue that any CPI measure of the implicit insurance service charge should include fraudulent claims as part of that charge. However, measuring the level of fraudulent claims is very difficult and any feasible pricing method may have to assume a constant level of fraudulent claims for practical purposes.

Concluding remarks: The "gross premiums, net expenditures" weighting approach used for insurance in the Canadian CPI seems more appropriate than "net premiums, gross expenditures" weights given the primary purpose of the index. The "gross premiums, gross expenditures" approach is a possible alternative weighting basis that may be adopted in Canada. Weights based on net premiums and gross expenditures would be more appropriate if the "implicit service charge" concept of insurance were adopted.

Price change for insurance is currently measured based on changes in gross premiums, holding all insurance policy provisions constant but not the risk probabilities or levels of payouts of claims. Thus the "security of the consumer" is the quality being held constant and any changes in risk factors reflected in premiums are reflected as price change. This seems appropriate as a reflection of the price to consumers of the insurance input to private transportation services. When vehicle insurance is viewed as a cost of owning and operating a private vehicle to obtain private transportation services (or tenants' insurance is viewed as a part of the cost of obtaining rental shelter services), any change in gross premiums is seen as a change in the price to consumers of the service purchased, whether the premium change arises from a change in the implicit service charge or in expected claims.

Changes in risk probabilities affect the level of expected claims. Therefore, if one wants to hold risk factors constant, it is conceptually equivalent to pricing only the "implicit service charge" of insurance. Given the diversity of the risk factors, their effects through time and the difficulties of measurement (even more difficult for the timely requirements of a non-revisable CPI), there is no practical direct method of pricing the change in insurance premiums holding all risk factors constant. The best apparent method of measuring the implicit service charge only that is conceptually defendable and practical might be some form of management services proxy.

This paper argues that even if a direct method of holding all risk factors constant could be developed, it would not necessarily be desirable for the purposes of the official Canadian CPI, based on a risk-assuming definition of insurance services. Tracking changes in the gross premiums of insurance policies of constant terms and conditions without adjusting for risk probabilities is the most appropriate method for a CPI whose primary purpose is to reflect changes in the purchasing power of the consumer dollar. .

Appendix 1: Ralph Turvey: Note on Insurance, extract from his website

Life insurance and pension fund contributions are not appropriately included in a consumer price index, but there seems to be general agreement that household expenditure on insurance which provides some reimbursement of expenditure on medical or emergency services or on the repair or replacement of damaged or stolen goods ought to be included. There appear to be two alternative internally consistent methods of doing this:

Net net. The weight should be payments less claims paid and less changes in actuarial reserves. This net amount equals insurance companies' costs and profits for the service they provide:- collecting premiums and paying claims. Similarly, the price indicator should relate only to net premiums. Household expenditure on cars, furniture, health services etc., which are paid for by insurance claims would then be dealt with in the same way as all other expenditure on these things.

Gross gross. The weight should be gross premium payments (the weights in the rest of the index for expenditure on health services, cars, furniture, etc., excluding the part paid for by insurance claims) The price indicator should correspondingly relate to gross premium payments.

(1) measures the service charge, the cost of the administrative functions and risk-pooling provided by insurance companies. A rise in repair and replacement costs, in health service prices or a rise in risks will not raise this component of the index unless accompanied by an increase in the companies' costs or profits.

(2) could measure the cost to consumers of buying a constant amount and quality of repairs, replacements, health services etc through the intermediation of insurance. A premium increase reflecting higher prices of these services as well as higher insurance company administrative costs and profits would raise the index, but a premium increase reflecting increased risks should not. On the other hand, looking at the issue purely from the consumers' point of view, it could measure the cost to them of preserving a certain level of security against the financial consequences of ill health, fire, accident and theft. In this case, not only a rise in health service prices or repair and replacement costs but also an increase in risks would raise the index, reflecting a greater cost to consumers of obtaining the level of security which they had in the reference period.

The choice made must depend upon the relative feasibility of these alternatives. Since this will vary between countries for institutional reasons, no generalisations can be offered. The main issues are as follows:

Is it easier (1) to establish weights which do not distinguish expenditure financed by insurance claims from other expenditure and to deduct claims paid from premiums to obtain a weight for net premiums, than it is (2) to establish weights reflecting gross premium payments and to exclude from expenditures on health services, cars, furniture, and so on, those which are financed by claim payments?

In (2), claims paid to reimburse consumers will have to be deducted from those of their expenditures which the claims helped to finance. In (1) it is the other way round; claims paid on consumers' behalf directly to the suppliers of repairs, replacements, medical treatment, and so on, will have to be added to their own expenditures. In (1), in order to obtain net premium payments from the gross payments recorded in a household survey, the share of gross premium payments in the reference period which were received back as claim payments or added to reserves will have to be ascertained from insurance companies or estimated from their published accounts.

In (2), gross premiums must be ascertained for a sample of policies holding constant such characteristics as deductibles and, maybe, (with more difficulty) risk

Gross premiums are fixed by insurance companies in the light of their risk expectations and the competitive situation. In any month some of the claims settled will be on policies taken out in a previous year and claims on some of the policies for which premiums are currently paid will not be settled until a future year. How can a monthly index of the cost of the administrative and risk-pooling service provided by insurance companies, as required by (1), be calculated?
The actual cost of the service in any year, if defined as that year's premiums less that year's claims paid, can be negative in years when claims exceed premiums. The difference is covered by a reduction in insurance companies' reserves. The companies fix premiums so that, together with interest earned on their reserves, they cover expected claims plus the expected cost of the service, including normal profits *on average over future years*. This problem might be resolved by dividing the year's total value of claims by the actual average number of claims per policy in the year and multiplying by the average number of claims per policy computed for the past five years.
An index of the cost of the insurance cover provided by insurance companies, as required by (2) requires the measurement of premiums for a sample of clearly specified types of insurance. The question then arises whether the gross premiums should be quality adjusted for changes in risks as well as for changes in the prices and repair costs of insured goods and altered health service prices. If not, correction would be limited to the changes in health service prices and the prices and repair costs of the goods insured, thus accepting any increase in premiums stemming from increased risks as a price increase. In this case the premiums on a sample of insurance policies for specified objects would each be deflated by an appropriate price sub-index. Alternatively, no correction would be made, in which case the quality being held "constant" would be the security of the consumer, rather than the volume of the repairs, replacement and health treatment paid for by claims.[12]

Faced with all these problems, most countries estimate a sub-index for *gross* premiums, unadjusted for changes in risk, but give it a *net* weight. Gross premiums are thus treated as a proxy for net premiums. This approach has been adopted in the European Harmonised Index of Consumer Prices, the text of the relevant Regulation being available at. europa.eu.int\eur-lex\en\lif\1999\en_399R1617.html

Adoption of this sensible practicable approach raises a problem when insurance premiums are subject to taxation levied on the gross premiums. Adoption of the sensible practicable net-weight, gross-premium approach raises a problem when the gross premiums are taxed. The solution is to decompose the net service charge weight into two components: the charge before tax and the tax. The first component then provides the weight for gross premiums before tax, and the second component provides the weight for the tax.

[12] Even an indicator of this latter type can be complicated. The article "The measurement of liability insurance premium trends and their inclusion in the French consumer price index" by A. Marret, reprinted in the appendix of the 1989 ILO Manual, explains how such an indicator can be calculated.

Adrian Ball and David Fenwick[13]: Costing mobile telephone calls: the use of constrained user profiles

Abstract

The traditional tariff approach to the costing of mobile telephone calls fails to take into account the full range of package options available to consumers over the life of the "fixed" basket. Selecting packages at random and applying them to the most appropriate user profile has the effect of missing important price changes and also suffers from a number of other measurement weaknesses. The analogous situation for non-services is where there are two identical tins of baked beans standing side by side on the shelf in the same shop and the price collector prices the most expensive although we know that in reality the customer will buy the cheapest. The method proposed in this paper samples profiles and then selects for pricing the appropriate package from each service provider. It assumes a rational consumer who has perfect knowledge but this assumption is more realistic than the alternative of the ill-informed laggard, particularly in telecommunications in the UK where unit cost information is readily available and where at points in time customers can transfer from one tariff to another free of charge."

Key words: fixed basket, services, tariffs, price change, constrained consumer profiles.

Background

Mobile telephone charges have become a significant part of household expenditure in the UK and have been incorporated into the Consumer Price Index since 1998.

The measurement of the prices charged for telecommunications services and the construction of an associated price index is fraught with difficulty because of two particular factors:

There can be two prices co-existing for the same service delivered by the same service provider. These two prices co-exist when a new tariff is introduced without the previous being withdrawn.
The opportunity to change tariff and service provider is often constrained by the existence of signed contracts that commit the buyer to a particular tariff and particular service provider over a minimum period.

These difficulties are in addition to the generic problems relating to the construction of price indices for services- the latter ranging from the definition of the service provided and the identification and valuation of any change in the service, to conceptual issues about the timing of delivery and the relative merits of the three main approaches to price measurement and which is most appropriate: acquisition, payments and user costs (or consumption)[14].

[13] Office For National Statistics, UK

[14]

- ***Acquisition –*** the total value of all goods and services delivered during a given period, whether or not they were wholly paid for during the period, is taken into account
- ***Payment*** - the total payments made for goods and services during a given period, whether or not they were delivered, is taken into account

The solution adopted until recently by the UK Office for National Statistics has been to take a “tariff” approach that prices those call charges that would be experienced by typical “users” on selected tariffs. A sample of major suppliers is drawn and a price index is calculated for each of these suppliers and an overall price index for mobile telephones is calculated by weighting together the individual price indices for the selected suppliers using revenue figures provided by the industry regulator. The individual supplier-specific price indices consist of weighted averages of separate indices for pre-pay and contract telephones, weighted according to the number of customers using the two-types of package in each of the individual companies (revenue figures are unavailable at this level). These separate indices are constructed by measuring the changing cost of the most “popular” package for each of five customer profiles[15].

The drawbacks of this approach are threefold:

The initial sampling of packages directly from each of the selected service providers in reality can be problematic because of difficulties in identifying the most appropriate package to price and it also severely restricts the coverage of the sample which then becomes over-reliant on the pricing of a relatively small selection of packages. This is particularly problematic when, as in the case of mobile telephone, there is a fast changing and exceptionally diverse range of packages on offer to the customer. Past experience has shown that the resulting price index can be relatively volatile and unrepresentative of the price changes observed more generally in the mobile telephone market.
The resulting index quickly becomes unrepresentative as it continues to price packages that in reality few customers use. It follows packages for as long as they are available and ignores, for instance, the fact that some packages may be closed to new customers. This also adds to lack of representiveness.
The resulting index doesn’t take into account the reduced prices people pay from exchanging one tariff for another even when this does not involve a transfer to a different service provider and does not entail a change of service.

These problems are further accentuated by the lack of detailed expenditure data on the different packages provided by each service provider. This forces the use of (un-weighted) elementary indices at a higher level than would be ideal for the accuracy of the index.

It is against this background that the ONS investigated alternative solutions.

There are two main alternatives to the tariff approach described above:

- ***User cost (or consumption)*** – the total value of all goods and services consumed during a given period is taken into account.

The distinction is particularly important for purchases financed by some form of credit, notably major durable goods and housing, which are acquired at a certain point of time, used over a considerable number of years, and paid for, at least partly, some time after they were acquired, possibly in a series of instalments. It also is an issue for services where payment and delivery do not coincide. The usual approach is to take the price at the time the right to use the service is acquired and the service is available under the terms of the contract of sale.

[15] Five different types of user profile are priced for both pre-pay and contract customers. The user profiles are based on research commissioned by the industry regulator OFTEL. These cover volume of usage- ranging from “low” to “high” volume- time of call (peak, off-peak, weekend), and destination (local, national, own network, other network),. Only one type of user profile is priced for a particular contract or prepay package; the allocation of a user profile to a particular package is done subjectively. The index for a particular package/user profile is derived from the changing cost of the average monthly bill; this is calculated taking account f line rentals, (for contract customers), calls (by time of day and destination) and text message. When two packages are priced for a particular supplier, the resulting indices are weighted together to form a single index for the supplier, using weights according to the proportion of subscribers with the relevant user profile.

The "average bill" method. This method tracks the change in the average bill, expressed in unit cost terms, of all or a representative sample of users. The latter is stratified as appropriate by user profiles to ensure a broad range of users. In effect the resulting index would reflect a) changes in usage and b) changes in service providers. Such an index is not consistent with the fixed basket approach underlying the RPI and HICP because it reflects changes in consumption as well as price. In practice it can also be ruled out in the UK because of lack of access to detailed data from the telephone companies and the resources that would be required to quality assure it.

The "user profile" method. This method tracks the change in the average bill, expressed in unit cost terms, for a selection of fixed profiles relating to use, which are then weighted together by expenditure shares to construct an overall price index. This method does not raise the same conceptual issues that are associated with the "average bill" method and the data requirements are less daunting.

It is the "user profile" method that ONS decided to explore in more detail. It essentially reverses the sampling procedures underlying the "tariff" approach by sampling user profiles and then selecting from each service provider the appropriate packages based on cost to the user.

The attractions of this approach from a practical viewpoint are threefold:

The availability of explicit expenditure weights for different users profiles that can be used to construct appropriately weighted indices. ONS has ready access to expenditure data from OFTEL, the regulatory authority, on 21 detailed user profiles.
Easier identification of the "appropriate" package from each supplier for each user profile – all other things being equal this is taken as the cheapest. The subjective element associated with the current method is eliminated and the new method is more transparent.
Broader coverage through the costing of each user profile for each supplier and a better balanced sample through the use of expenditure information.

The "user profile" method also has a number of other advantages relating to fundamental issues of concept and measurement:

Product definition. The measurement of service prices offers a particular challenge because of the difficulty in identifying discrete service elements. For example, in the case of mobile phone services, is the purchase most appropriately defined in terms of the individual package available from a particular company or in terms of the number of call minutes purchased? The methodology previously used by ONS took as its starting point the former and sampled mobile phone packages much in the same way as for goods. But from an alternative viewpoint, customers can be said to be buying calls and the text messages that they are making, with prices set by the packages, rather than buying packages as such. A logical conclusion of this approach is that the price index should take account of consumer substitution between packages, particular where this occurs within the same service provider.

Fixed basket concept. If customers can be said to be buying calls and the text messages that they are making, with prices set by the packages then it follows that under the terms of a fixed basket it is the user profiles that should remain fixed throughout the year

The logical conclusion is that the construction of the price index should be based on the tracking of the unit cost of calls for fixed customer profiles but allowing some substitution between packages.

In following this approach practical considerations arise in relation to the sampling of the packages used to price each user profile and the circumstances under which substitution between packages should be incorporated:

Sampling of packages. It can be argued that as customers will normally be expected to select the cheapest package for their anticipated usage it should be the latter that provides the basis for the pricing of the initial package at the point of chain linking. Whether in practice this is done across the whole market or separately within each service provider depends on the availability of data, although clearly there is a preference for the latter as this effectively increases the sample size by re-computing prices for each customer profile across each of the individual service providers. The latter is the approach which is available and has been adopted in the UK- detailed tariffs and explicit expenditure weights for the four mobile telephone companies (Vodaphone, O2, Orange and T-Mobile) facilitates the calculation of individual company indices and weighted averages to produce an overall index. Thus, the most economical package was priced within each company for each customer profile.

Substitution between packages. Clearly account needs to be taken of any restrictions on the movement between packages imposed by the type of contract a customer has with the mobile phone service provider. In the UK there are two main types of contract:

Pay as You go. Under this arrangement the customer pays for the handset. The latter is locked to the network of a particular service provider, usually at a discount. Other than that there is no formal contract with the service provider and customers are free to move from one package to another package offered by the same service provider without a financial penalty, though changing provider will cost money as the customer will need to buy a new handset. There is no line rental but call costs are higher.
Monthly contract. Under this arrangement there is a formal contract with the service provider- normally for a minimum 12 month period- where a handset is provided free of charge and a monthly fee is paid to cover line rental and some call costs. In general call costs are lower than for Pay as You go to reflect the forward commitment.

Clearly different computations will need to carried out to reflect the different opportunities customers have for moving from one package to another as dictated by the particular circumstances and constraints associated with each type of contract. In particular, "Pay as you Go" customers are free to change without financial penalty between packages within a service provider (usually by making a phone call to request a change) whilst monthly contract customers will face a financial penalty if they switch service provider before the end of the 12 month minimum contract period. These differences were addressed by specifying different rules for computation as follows:

Pay as You Go

Users are free to change package as they wish;
It is costly to swap between service providers, as this would involve the purchase of a new handset.

In this case, taking a rational consumer as the basis for monthly comparisons, it would seem appropriate to allow each user profile to switch freely between packages, picking up the cheapest available each month. However, this movement should be restricted to the same service provider as in reality customers are unable to move service providers without financial penalty. Therefore, for each profile, we select the cheapest package available for each service provider in each month, and calculate an index for each combination. These are then weighted together over the profile to give a PAYG index for each company.

Monthly contract

In the UK market it is more difficult for monthly contract customers to move between packages, even within the same service provider.
This is because the cost of the handset provided is largely offset against the monthly bill, and the companies require the user to be on a contract for a year to ensure that they recoup the cost of the hardware.

In these circumstances it is not appropriate to allow all users to move to the lowest package available to them each month.

The appropriate approach is to select the cheapest package available to each user profile at the point of chain linking in January. Users are then divided into 12 cohorts reflecting the proportions that we would expect would be due to renew contracts in each of the following 12 months. The 1st cohort is then assumed to look for a new cheapest package in February, simulating those whose 12 month contract comes to an end in that month. The other groups continue on the package selected in January, as they are still looked into it. In March, it is assumed that the second cohort selects a new cheapest package, with the first cohort continuing on the package selected in February, and the rest on the package selected in January.

This continues throughout the year, with one cohort re-selecting the cheapest package each month, and the rest carrying on with the package they were on in the previous month.

As for PAYG, the indices produced are weighted together to produce an overall monthly contract for each company.

Final Indices

Price indices and then computed for each service provider by weighting the two contract types together by sales weights. The final price index for all service providers is a weighted average of the individual indices for each service provider, with weights based on expenditure shares.

Keith Woolford:[16] Experimental Price Indexes for Financial Services

For some years now there has been considerable community interest in the prices paid by households for financial services. An outcome of a review of the Consumer Price Index (CPI) undertaken in 1997 was a commitment by the Australian Bureau of Statistics (ABS) to develop a price index for financial services for eventual inclusion in the CPI. This index was to be designed to cover the explicit fees and charges paid by households as well as any indirect costs embodied in interest rate margins.
This paper presents an experimental index measuring price changes for some of the most significant financial services acquired by households — deposit and loan facilities provided by financial institutions and services associated with the acquisition and disposal of shares and real estate. The index also covers taxes and stamp duties levied on the relevant transactions. It represents the first attempt by a national statistical agency to construct such a comprehensive measure of price change for financial services.

The paper has been published as ABS information paper 6413.0, accessible through ABS Web-site: http://www.abs.gov.au/AUSSTATS/abs@.nsf/papersbyTopic

[16] Australian Bureau of Statistics

Session 2: Relationships between consumer price indices and other statistical systems

Moderator: Bert M. Balk, Statistics Netherlands

Summary of Session

In this session two papers were presented. The first one, entitled 'CPI and Household Income Expenditure under Deflationary Trend' and written by Noboyuki Sakashita and Makoto Shimizu, shows that the common assumption of a representative household with homothetic preferences is empirically unjustified. There is ample evidence that CPI's and subindices thereof for household groups defined by income level show differential development.

The issue of consistency-in-aggregation (CIA) quite naturally comes up in economic accounting systems such as the National Accounts. Nominal values are by nature additive. Likewise, one wants real values (that is, deflated nominal values) to be additive. This implies a mathematical relation between aggregate and subaggregate price and quantity indices. The same sort of relation emerges when the CPI is calculated stepwise and quite natural consistency requirements are posed. Both relations appear to be a specific case of CIA, as formally defined and discussed by Balk (1995).[17] It was shown that requiring CIA in addition to a few other, quite natural requirements effectively reduces the admissible indices to those of Laspeyres and Paasche.

The paper presented by Heikki Pursiainen, 'The Algebraic Interpretation of Consistency in Aggregation and Quasilinear Index Numbers', provides an in-depth treatment of the issue of CIA, and specifically proposes a generalisation of the definition of Balk (1995). It is, however, reassuring to see that, when restricted to price and quantity indices, Pursiainen's definition reduces to Balk's definition.

Recommendations for Statistical Agencies

Consideration should be given to CPI's for household groups, *e.g.* defined by income level.
Consideration should be given to a CPI that exhibits CIA next to a CPI that does not exhibit CIA but has a number of other desirable properties (such as superlativity).

[17] Balk, B. M., 1995, "Axiomatic Price Index Theory: A Survey", *International Statistical Review* 63, 69-93.

Nobuyuki Sakashita18and Makoto Shimizu[19]: CPI and Household Income and Expenditure under Deflationary Trend[20]

Abstract

General index of consumer price in Japan fell 0.3% in the year 2003 from the preceding year, showing considerable reduction in declining rate compared with 0.9% in the former year. This result, however, reflects rise in public utilities and does not necessarily reflect supply and demand for commodities and services. Thus, the government and the central bank share recognition that deflation is still continuing, although decline in consumer prices has reduced.

On the other hand, the results by yearly income groups show that declining rate for the first yearly income quintile group[21] with the smallest yearly income is 0.2 percent point smaller than that for the fifth income quintile group with the largest yearly income in the year 2003. This may indicate that income structure changes under the recent deflationary trend.

Examining the results by characteristics of items, index for basic expenditure items, whose expenditure elasticity are less than 1, fell by only 0.1%, while that for selective expenditure items fell by 0.6%. Additionally, index for items in annual purchase frequency classes of less than 0.5 times and over fell by 1.7%, while that of 15 times or more rose by 1.5%.

In argument about the cost-of-living index, in order to compensate for the “upward bias” of the Laspeyres-type fixed basket index, alternative formulae as the geometric mean or the CES type index are proposed, whose models are based on the assumption that a representative household with average income, average expenditure and homothetic preferences exists, in other words all items are common goods having constant elasticity of substitution.

In reality, superior goods and inferior goods coexist in the market and their optimum allocations may change according to incomes.

Discrepancy from homothetic model may emerge and effect differently among social strata especially in the year 2003. The Family Income and Expenditure Survey in Japan supplies us with tools to analyze such issues every month. Here we try, rather than considering theoretical verifications, to examine changes in prices and household expenditures by social strata, and to compare price indices in some formulae.

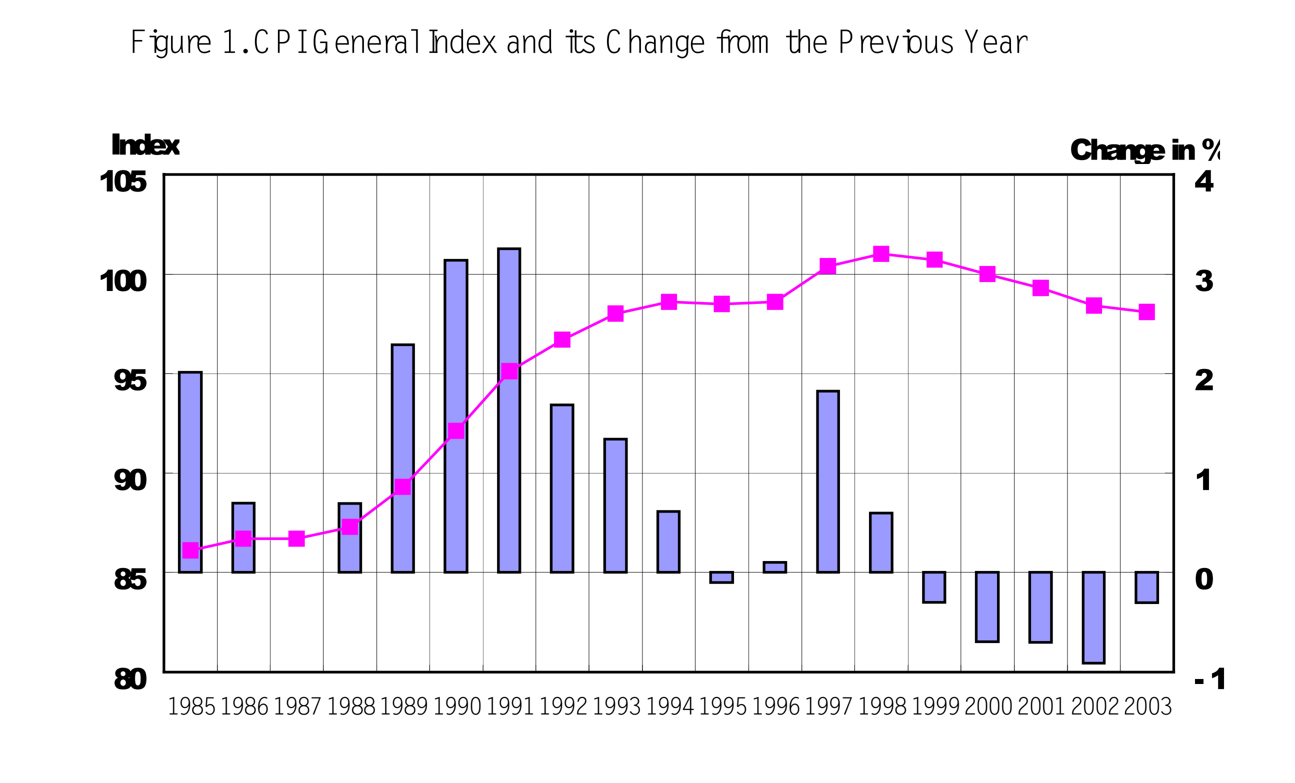
1. Outline of Recent CPI in Japan

General index of consumer prices in Japan attained the peak 101.0 in 1998 (2000=100) and stood at 98.1 in 2003, fell 0.3% from the preceding year, recording the fifth year of consecutive decline (Figure 1). Japan has never experienced such a long decline of CPI after World War II. But CPI in 2003 shows considerable reduction in declining rate compared with 0.9% in the former year.

CPI moves quite differently among goods and services. Figure 2 shows changes of CPI during five years, crossing data segmented by the peak year 1998; between in 1993-1998 when CPI rose and in 1998-2003 when CPI declined by ten major groups. Changes in 1998-2003 are lower than in 1993-

[18] Science Council of Japan
[19] Statistics Bureau, Japan
[20] Opinions expressed in this paper are those of the authors and do not necessarily represent official views of the Statistics Bureau.
In this paper, weights from the Family Income and Expenditure Survey are applied almost in the same way as in the official indices but without detailed arrangements.
[21] Yearly income quintile groups mean the five equally divided groups in terms of the number of households, after arranging them in order of the yearly income.

1998 in nine major groups except for in Transportation & communication. In both periods, CPI rose in Education, Medical care, Housing and Miscellaneous[22], declining in Transportation & communication and Furniture & household utensils. In other four groups including basic consumption as Food, Fuel, light & water charges and Clothes & footwear, CPI rose in 1993-98 and dropped in 1998-2003. This may have reduced differences between the rich and the poor.



Indeed, expenditures differ among income groups. The richest group; the fifth group by yearly Income Quintile Groups spends a lot for Transportation & communication and Reading & recreation as well as Food, while the first group, which consists of those with small yearly income does not spend a lot for those except for Food (Figure 3). Expenditure for Education regarded as typical selective expenditure in the first group is about one eighth for the fifth group in amount, or almost half in weight, and staying low partly because of the price rise even during the deflation period. On the other hand, expenditures for Medical care, Housing and Fuel, light & water charges, regarded as basic expenditures, are similar between these two groups. Comparatively low level of price changes except for Education seems to have reduced the gap as to expenditures between the first and the fifth group. Ratio of total consumption expenditure of the fifth quintile group to the first quintile group tends to drop for recent years (Figure 4). Taking other observations into consideration[23], the gap of expenditures between the rich and the poor seems to have shrunk in recent years when price hikes have not be seen for most of commodities and services.

[22] Personal care services, Toilet articles, Cigarettes, etc.
[23] For example, ratio of total consumption expenditure of the fourth quintile group to the second quintile group recorded sixth year of consecutive decline in 2003.

Figure2. Change during Five Years by Ten Major Group (%)

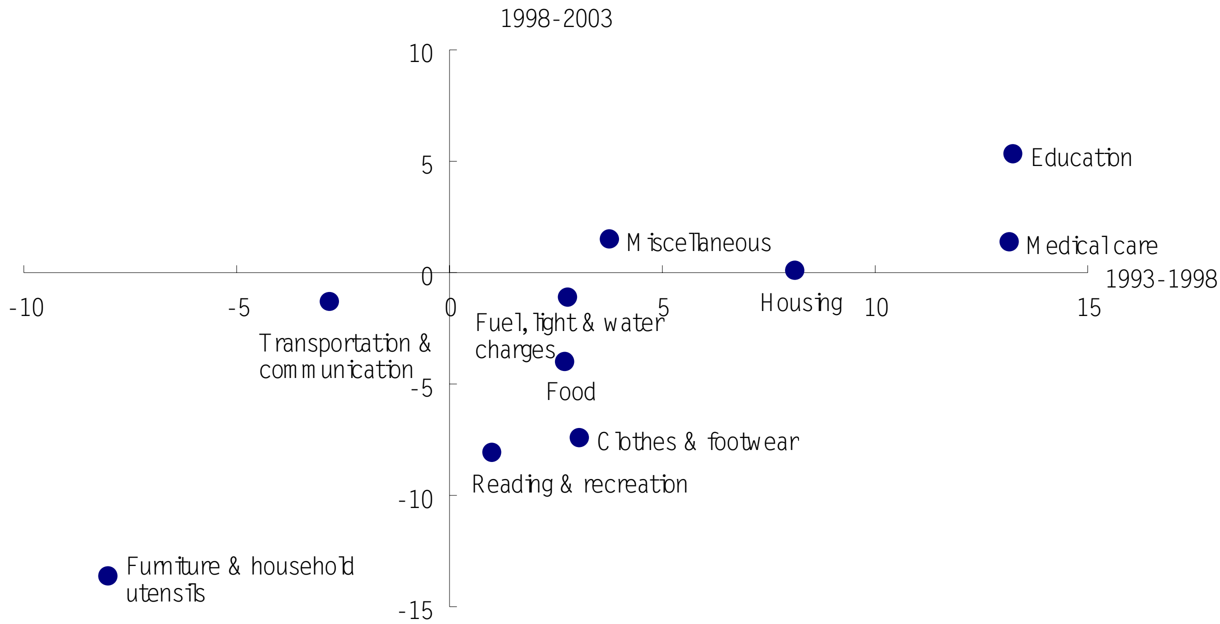


Figure 3. Expenditures by Income Quintile Group for the Year 2003

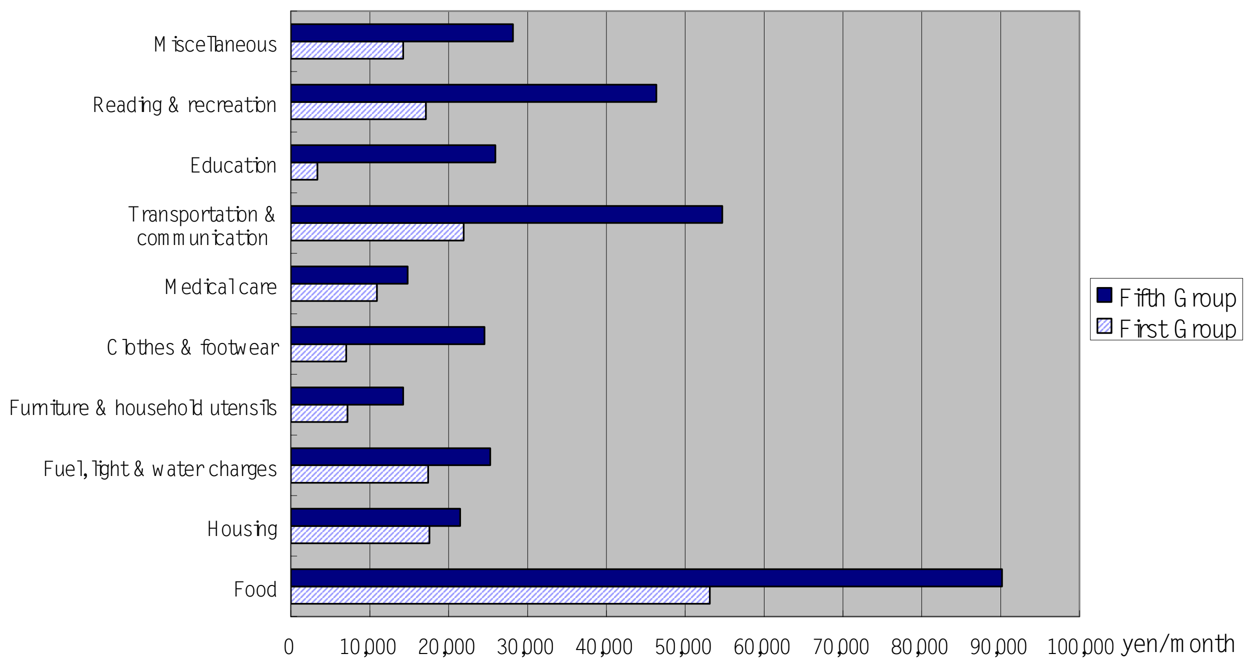
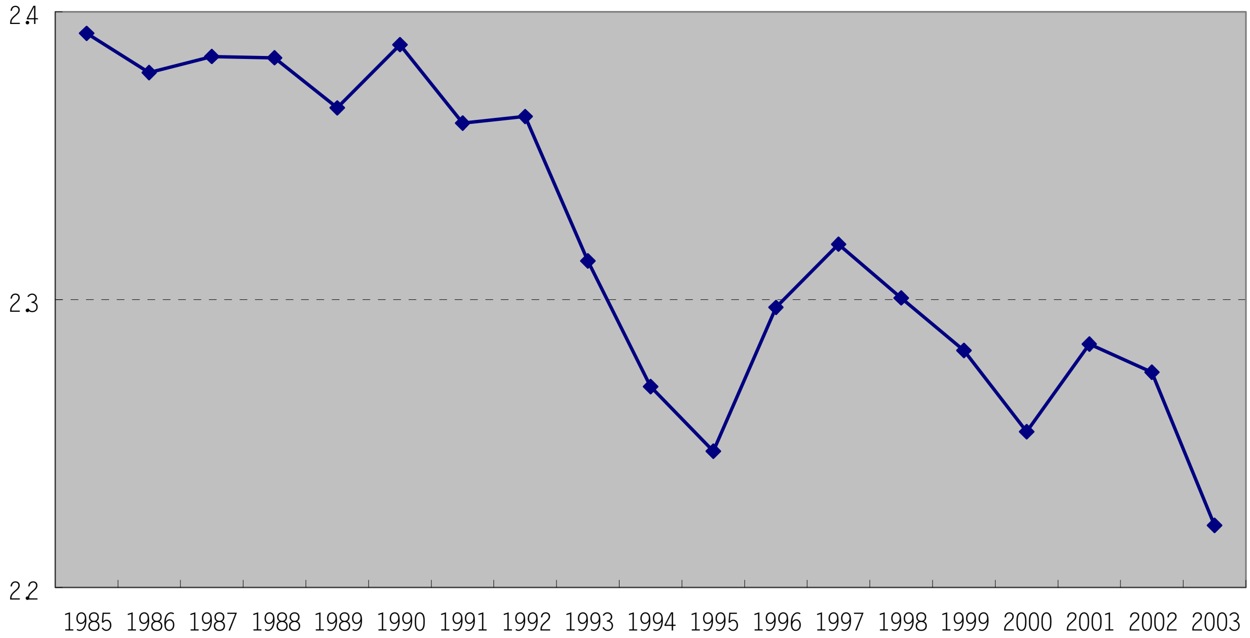


Figure 4. Ratio of Total Consumption Expenditure of the Fifth Quintile Group to the First Quintile Group



2. Characteristics of CPI in the year 2003

Considering sub-indices by goods and services classification in the year 2003, goods index fell 0.8% mainly due to fall in industrial products such as durable goods and food products (Table 1). While services index rose 0.4% mainly due to 1.3% rise in public services such as medical treatment. This means decline of fall of CPI in 2003 reflects rise in cost for public utilities and does not necessarily reflect supply and demand of commodities and services in the market.

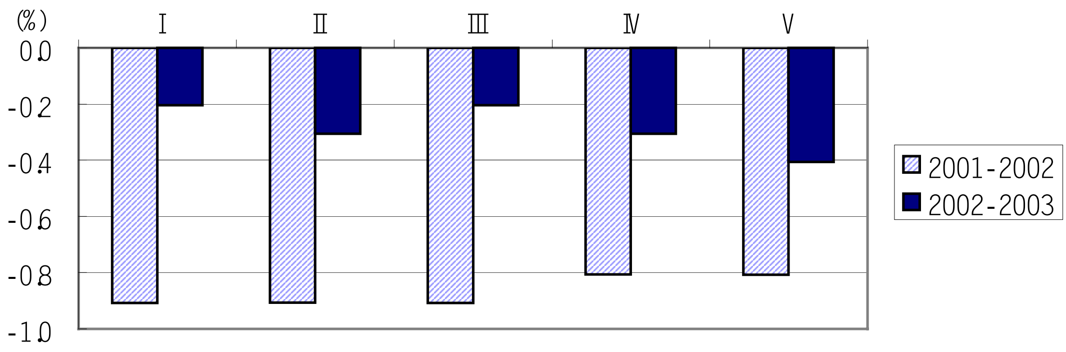
Table 1. Change of CPI from the Previous Year of Goods and Services

	2001 %	2002 %	2003 %	Contribution (2003)
Goods	-1.4	-1.8	-0.8	-0.41
Services	-0.1	0.0	0.4	0.20
Public Services	-0.2	-0.1	1.3	0.17
General Services	0.0	0.0	0.0	0.00

On the other hand, examining the results by yearly income groups[24], declining rate for the first yearly income quintile group is now 0.2 percent point smaller than that for the fifth income quintile group (Figure 5). This may indicate that income structure changes under the recent deflationary trend.

[24] In chapter 2, 3 and 4 in this paper, target population for weight by yearly income quintile group excludes agricultural, forestry and fisheries households.

Figure 5. Change of CPI from the Pevious Year by Yearly Income Quintile Group



Examining the results by characteristics of items, index for basic expenditure items, whose expenditure elasticity are less than 1, fell by only 0.1%, while that for selective expenditures fell by 0.6% (Figure 6). Additionally, index for items in annual purchase frequency classes of "less than 0.5 times" fell by 1.7%, while that of "15 times or more" rose by 1.5% (Figure 7). The difference of price changes between essential expenditures and nonessential expenditures may have affected structures of expenditures by income groups.

Figure 6. Change of CPI from the Previous Year by Basic and Selective Expenditure

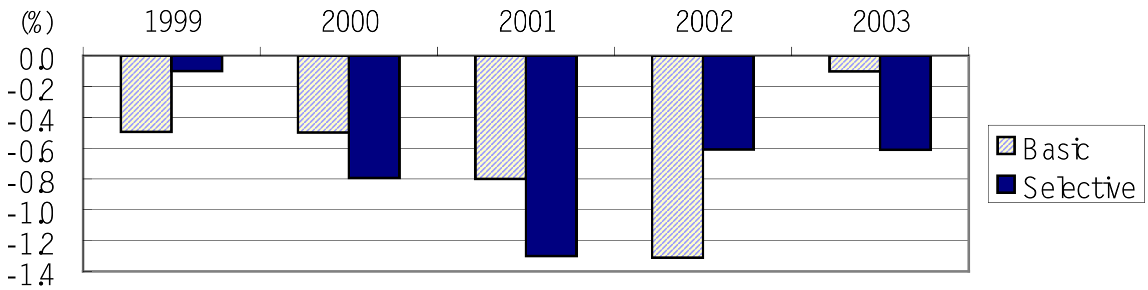
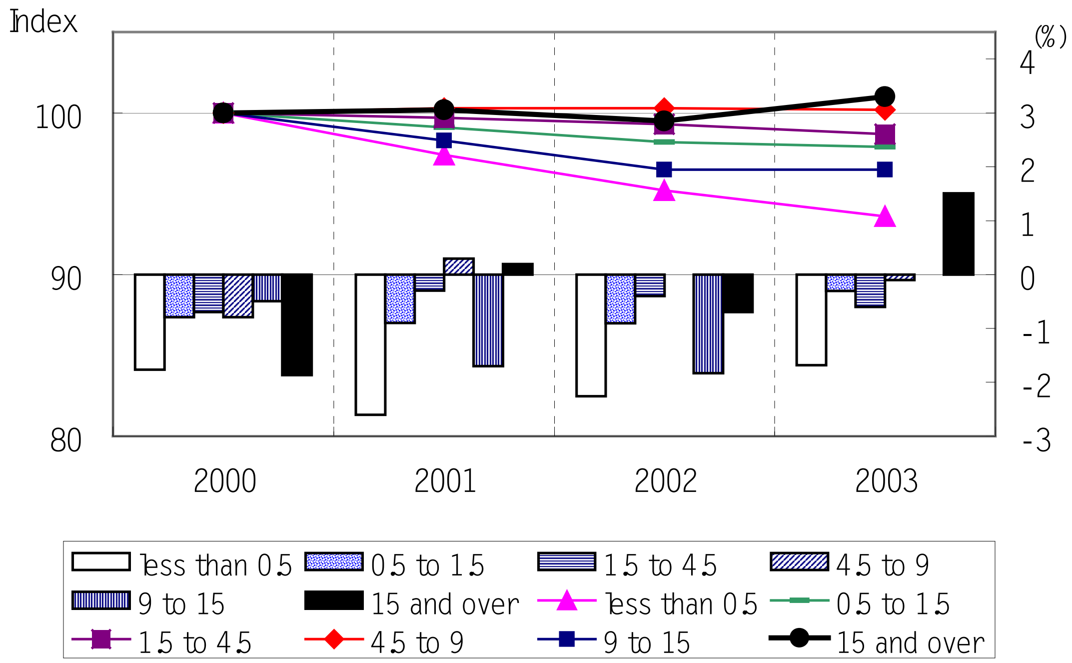


Figure 7. Price Index and Its Change from the Previous Year by Annual Purchase Frequency



3. Comparison among Types of Indices

Dividing households into yearly income quintile groups, we compare the Laspeyres-type index with the Paasche-type index for each quintile group as yearly average index for all Japan. The result is shown in table 2.1. The fact that discrepancy between the Laspeyres-type index and the Paasche-type Index is comparatively large in the group V, with large selective expenditures, they can enjoy benefits of decline in prices. The group with the smallest discrepancy is, however, the group II – not the group I – with more than 0.4 percent point smaller than the group V, while the discrepancy of the group I is larger than average, being closer to that of the group V.

Table 2.1. Laspeyres and Paasche Indices by Yearly Income Quintile Group for the Year 2003 (2000□100)

	All	I	II	III	IV	V
Laspeyres	97.8	98.0	97.9	97.8	97.8	97.6
Paasche	97.3	97.4	97.7	97.2	97.3	96.9
(P-L)×100/L	-0.5	-0.6	-0.2	-0.6	-0.5	-0.7

As characteristics in each year after the year 2000, Tables 2.2 to 2.4 show the Laspeyres-type indices and the Paasche-type indices for each year by income groups based on and in reference to the previous year as 100.

Table 2.2. Laspeyres and Paasche Indices by Yearly Income Quintile Group for the Year 2001 (2000□100)

	All	I	II	III	IV	V

Laspeyres	99.1	99.2	99.2	99.1	99.1	99.0
Paasche	98.9	99.1	99.1	98.9	98.9	98.8
(P-L)×100/L	-0.2	-0.1	-0.1	-0.2	-0.2	-0.2

Table 2.3. Laspeyres and Paasche Indices by Yearly Income Quintile Group for the Year 2002 (2001□100)

	All	I	II	III	IV	V
Laspeyres	98.9	99.0	99.0	98.9	98.9	98.9
Paasche	98.9	99.0	99.0	98.8	98.9	98.8
(P-L)×100/L	-0.0	-0.0	-0.0	-0.1	0.0	-0.0

Table2.4. Laspeyres and Paasche Indices by Yearly Income Quintile Group for the Year 2003 (2002□100)

	All	I	II	III	IV	V
Laspeyres	99.6	99.7	99.7	99.6	99.6	99.5
Paasche	99.6	99.6	99.7	99.6	99.6	99.5
(P-L)×100/L	-0.0	-0.1	0.0	0.0	-0.0	-0.0

The results show that from the year 2000 to 2001, in line with the popular theory, the Paasche-type indices are lower than the Laspeyres-type indices. And discrepancies for lower-income groups tend to be smaller than those for higher-income groups, because, in general terms, lower-income groups have tight supplies and demands, and they benefit less from substitution.

Small discrepancies from the year 2001 may underpin small movements of prices and household consumption. Meanwhile, The group IV in 2001 to 2002 and the group II and III in 2002 to 2003 show, however slightly, reversal of the Laspeyres-type and the Paasche-type indices. Reversal of these indices in not-so-small income group is against the common view described above.

Since this phenomenon may affect approximation of the cost-of-living index under certain assumptions for substitution, we compare the Laspeyres-type, the Paasche-type, the Fisher-type and the geometric mean indices since the year 2000 for each income group (Figures 8.0.1-8.5.2).

Figure 8.0.1. CPI - All Groups

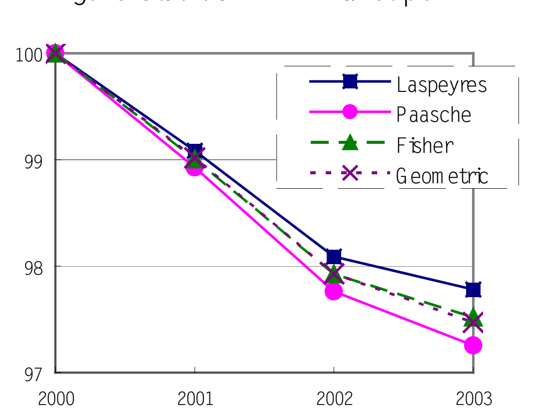
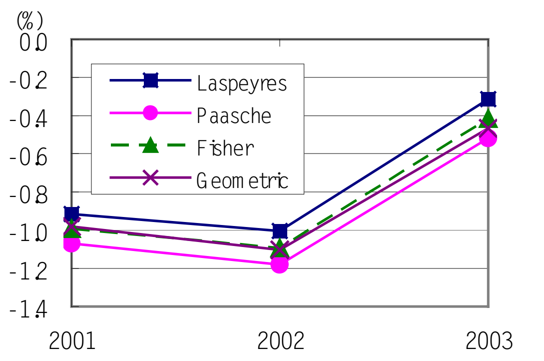


Figure 8.0.2. Change of CPI from the Previous Year - All Groups



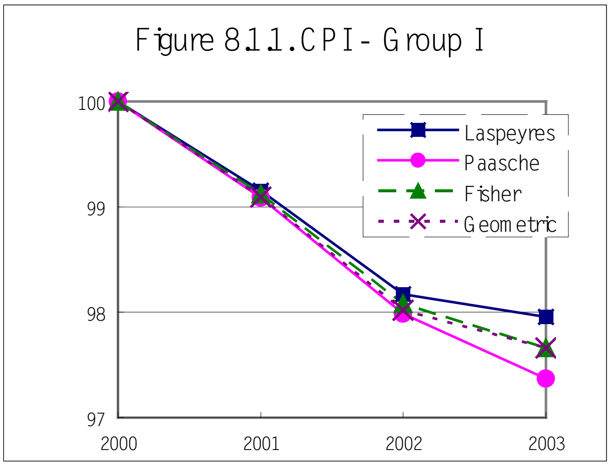


Figure 8.1.1. CPI - Group I
100
99
98
97
2000
2001
2002
2003
Laspeyres
Paasche
Fisher
Geometric

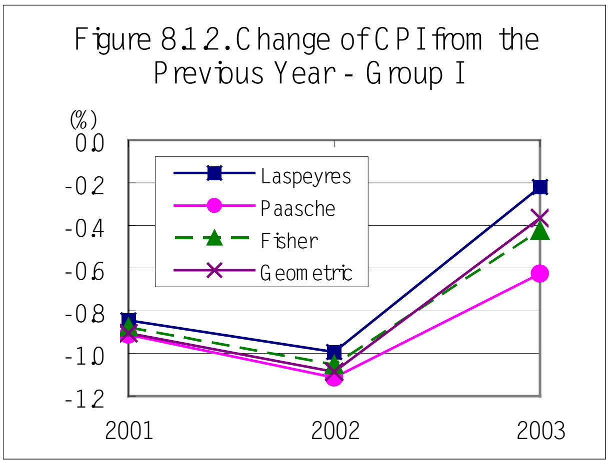


Figure 8.1.2. Change of CPI from the Previous Year - Group I
(%)
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
-1.2
2001
2002
2003
Laspeyres
Paasche
Fisher
Geometric

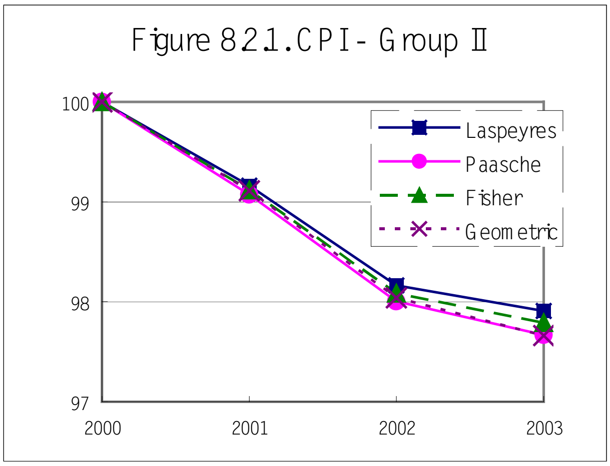


Figure 8.2.1. CPI - Group II
100
99
98
97
2000
2001
2002
2003
Laspeyres
Paasche
Fisher
Geometric

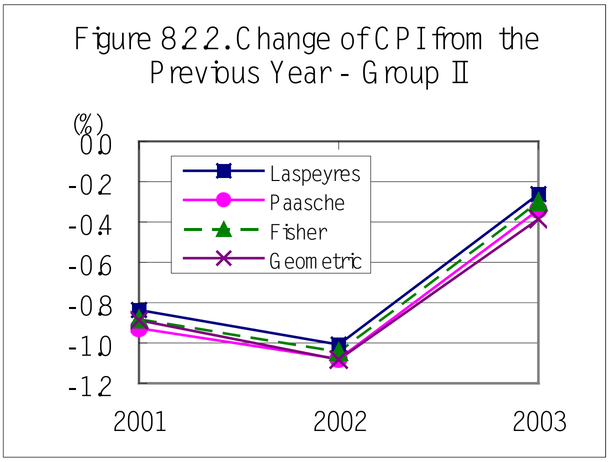


Figure 8.2.2. Change of CPI from the Previous Year - Group II
(%)
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
-1.2
2001
2002
2003
Laspeyres
Paasche
Fisher
Geometric

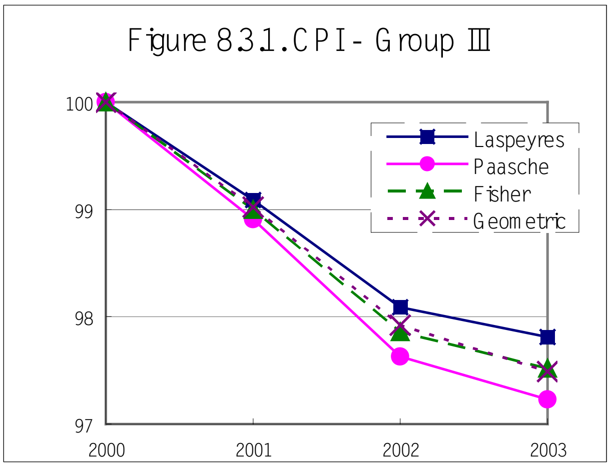


Figure 8.3.1. CPI - Group III
100
99
98
97
2000
2001
2002
2003
Laspeyres
Paasche
Fisher
Geometric

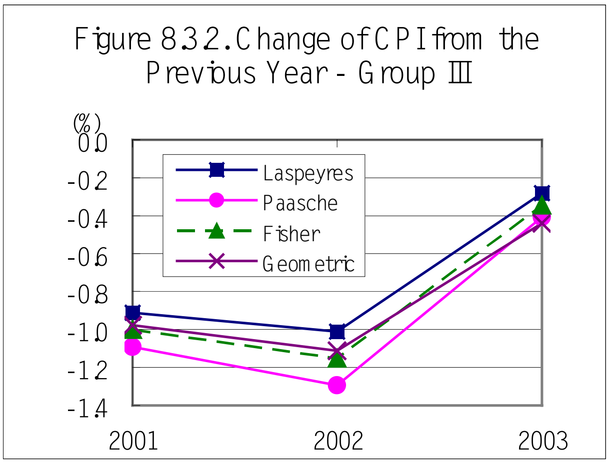
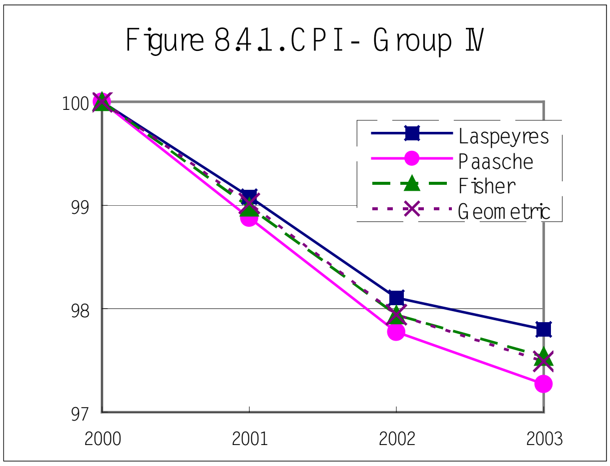
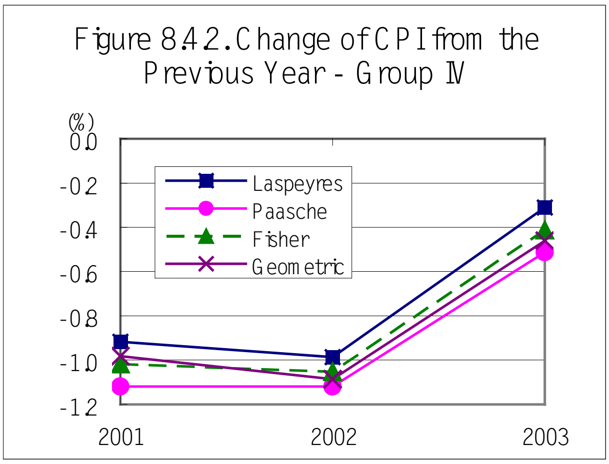
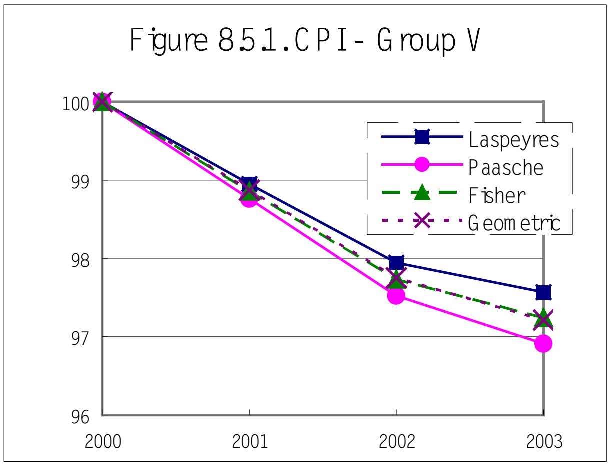
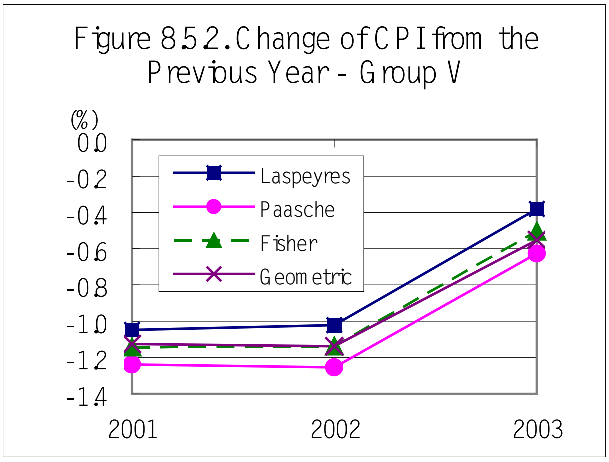


Figure 8.3.2. Change of CPI from the Previous Year - Group III
(%)
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
-1.2
-1.4
2001
2002
2003
Laspeyres
Paasche
Fisher
Geometric









These results show that in indices for all, the geometric mean indices show quite similar movements to those of the Fisher-type indices (the ideal formula), located between the Laspeyres-type indices and the Paasche-type indices. In indices for each group, however, the geometric mean indices sometimes show discrepancies from the Fisher-type indices, and in the year 2003, changes from the previous year of the geometric mean indices are lower than those of the Paasche-type indices for the group II and the group III and the level of the geometric mean index is lower than the Paasche-type indices for the group II. This phenomenon is easily understood considering that the geometric mean indices are designed to show good performance with high elasticity of substitution, that is, in case prices and quantity fluctuate in the same directions, they may weaken adequacy.

4. Contribution Factors

Table 3 shows contributions for each subgroup of items to both the Laspeyres-type and the Paasche-type index of the year 2003 based on the year 2002 (shown in table 2.4) and their discrepancies to measure contribution by each subgroup to overall discrepancy as to the two types of indices. If price and quantity move in the opposite direction, as in usual substitution behavior, the contribution of the subgroup is negative, while if they move in the same direction, as in special goods like the Giffen good, the contribution is positive. The sum of the contributions is equal to the overall discrepancy. Comparatively high contributions mark in Medical service in the group II, III and V, and Recreational durables in the group III.

Medical service, whose price rose by the hike of self-pay ratio for health insurance in April 2003, may be a basic and significant service to households for which no other products and services can substitute. And, in this field, change of the target population as well as change of prices influence on macro-level purchase behavior.

As for Recreational durables, contributions differ significantly among income groups. This subgroup includes PCs with rapid quality improvement and price down. Demand for these products also depends on their diffusion in each social group.

Table 3. Contribution to the Annual Changes by Yearly Income Quintile Group of Laspeyres and Paasche Indices □2002-2003□

	Laspeyres						□aasche						□aasche-Laspeyres					
	All	□	□	□	□	□	All	□	□	□	□	□	All	□	□	□	□	□
General	-0.41	-0.28	-0.32	-0.44	-0.39	-0.50	-0.43	-0.37	-0.29	-0.43	-0.43	-0.51	-0.02	-0.09	0.02	0.01	-0.04	-0.01
Cereals	0.04	0.04	0.03	0.04	0.03	0.03	0.03	0.04	0.04	0.03	0.03	0.03	-0.00	-0.00	0.00	-0.00	0.00	0.00
Fresh fish & shellfish	-0.04	-0.04	-0.04	-0.05	-0.05	-0.04	-0.04	-0.04	-0.04	-0.04	-0.05	-0.04	0.00	0.00	0.00	0.01	0.00	0.00
Meat	0.02	0.02	0.01	0.02	0.02	0.02	0.02	0.02	0.01	0.02	0.02	0.02	-0.00	-0.00	0.00	-0.00	-0.00	-0.00
Dairy products & eggs	-0.01	-0.02	-0.02	-0.01	-0.01	-0.01	-0.01	-0.02	-0.02	-0.01	-0.01	-0.01	0.00	0.00	0.00	0.00	0.00	0.00
Fresh vegetables	0.08	0.09	0.08	0.09	0.08	0.07	0.08	0.09	0.08	0.08	0.08	0.07	-0.00	-0.00	-0.00	-0.01	-0.00	-0.00
Fresh fruits	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
Oils, fats & seasonings	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.03	-0.02	-0.00	-0.00	-0.00	0.00	0.00	0.00
Cakes & candies	-0.01	-0.01	-0.00	-0.01	-0.00	-0.01	-0.01	-0.01	-0.00	-0.01	-0.00	-0.01	0.00	0.00	0.00	0.00	-0.00	-0.00
Cooked food	-0.02	-0.03	-0.02	-0.02	-0.02	-0.02	-0.02	-0.03	-0.02	-0.02	-0.02	-0.02	-0.00	-0.00	-0.00	0.00	-0.00	-0.00
Beverages	-0.04	-0.05	-0.04	-0.04	-0.04	-0.03	-0.04	-0.05	-0.04	-0.04	-0.04	-0.03	-0.00	0.00	-0.00	0.00	-0.00	-0.00
Alcoholic beverages	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	0.00	-0.00	0.00	0.00	0.00	-0.00
Eating out	-0.01	-0.00	-0.01	-0.01	-0.01	-0.01	-0.01	-0.00	-0.01	-0.01	-0.01	-0.01	0.00	0.00	0.00	0.00	0.00	0.00
Rent, excluding imputed rent	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Repairs & maintenance	-0.02	-0.01	-0.02	-0.02	-0.03	-0.04	-0.03	-0.02	-0.01	-0.02	-0.02	-0.05	-0.00	-0.00	0.01	-0.01	0.00	-0.01
Electricity	-0.07	-0.08	-0.07	-0.07	-0.07	-0.07	-0.07	-0.08	-0.07	-0.07	-0.07	-0.06	0.00	-0.00	-0.00	0.00	-0.00	0.00
Gas	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.00	-0.00	0.00
Other fuel & light	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.03	0.02	0.02	0.02	0.02	-0.00	-0.00	0.00	-0.00	-0.00	-0.00
Water & sewerage charges	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-	0.0	-	0.0	0.0

	1	1	1	1	1	1	1	1	1	1	1	1	0	0.00	0	0.00	0	0
Household durables	-0.10	-0.08	-0.10	-0.11	-0.09	-0.11	-0.11	-0.09	-0.10	-0.12	-0.11	-0.11	-0.00	-0.01	0.01	-0.01	-0.02	0.00
Interior furnishings	-0.01	-0.01	-0.01	-0.01	-0.02	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.02	0.00	0.00	0.00	-0.00	0.00	-0.00
Bedding	-0.01	-0.01	-0.00	-0.01	-0.01	-0.01	-0.01	-0.01	-0.00	-0.01	-0.01	-0.01	-0.00	-0.00	0.00	-0.00	-0.00	0.00
Domestic utensils	-0.01	-0.02	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	0.00	0.00	0.00	0.00	0.00	0.00
Domestic non-durables	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.00	-0.00	0.00	0.00
Domestic services	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.00	0.00	-0.00	-0.00	-0.00	-0.00
Clothes	-0.06	-0.05	-0.05	-0.06	-0.06	-0.08	-0.06	-0.05	-0.06	-0.06	-0.07	-0.08	-0.00	0.00	-0.00	-0.00	-0.00	0.00
Shirts, sweaters & underwear	-0.03	-0.03	-0.03	-0.03	-0.04	-0.04	-0.03	-0.02	-0.03	-0.03	-0.04	-0.04	-0.00	0.00	-0.00	-0.00	-0.00	-0.00
Footwear	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	0.00	0.00	0.00	0.00	0.00	0.00
Cloth & other clothing	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	0.00	-0.00	0.00	0.00	0.00	0.00
Medicines & health fortification	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
Medical supplies & appliances	-0.02	-0.04	-0.03	-0.02	-0.02	-0.02	-0.03	-0.04	-0.03	-0.03	-0.02	-0.02	-0.00	-0.01	-0.00	-0.00	0.00	-0.00
Medical services	0.17	0.19	0.19	0.17	0.18	0.14	0.18	0.17	0.20	0.19	0.16	0.17	0.01	-0.01	0.01	0.02	-0.01	0.03
Public transportation	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.00	-0.00	0.00	-0.00	-0.00	0.00
Private transportation	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Communication	-0.01	-0.01	-0.01	-0.01	-0.00	-0.00	-0.01	-0.01	-0.01	-0.01	-0.01	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
School fees	0.04	0.03	0.04	0.05	0.06	0.05	0.05	0.03	0.04	0.04	0.06	0.06	0.00	0.00	0.00	-0.00	0.00	0.00
School textbooks & reference books	-0.00	0.00	0.00	0.00	-0.00	-0.00	-0.00	0.00	0.00	0.00	-0.00	-0.00	0.00	-0.00	-0.00	-0.00	0.00	0.00
Tutorial fees	-0.00	-0.00	-0.00	-0.00	-0.01	-0.01	-0.00	-0.00	-0.00	-0.00	-0.01	-0.01	0.00	-0.00	-0.00	0.00	-0.00	0.00
Recreational durables	-0.2	-0.1	-0.1	-0.2	-0.2	-0.2	-0.2	-0.2	-0.1	-0.2	-0.2	-0.2	-0.0	-0.0	0.00	0.02	-0.0	-0.0

	1	5	6	4	3	5	3	0	6	2	4	8	1	6			1	2
Recreational goods	-0.06	-0.05	-0.05	-0.06	-0.06	-0.05	-0.06	-0.05	-0.06	-0.06	-0.05	-0.05	0.00	-0.00	-0.00	0.00	0.00	0.00
Books & other reading materials	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	-0.00	-0.00	0.00	-0.00	-0.00	-0.00
Recreational services	-0.01	0.00	0.00	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Personal care services	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Toilet articles	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.03	-0.00	0.00	0.00	0.00	-0.00	-0.00
Personal effects	0.03	0.02	0.02	0.03	0.03	0.04	0.03	0.02	0.02	0.03	0.03	0.04	-0.00	-0.00	0.00	0.00	-0.00	-0.01
Cigarettes	0.02	0.03	0.02	0.02	0.01	0.01	0.02	0.03	0.02	0.02	0.01	0.01	-0.00	0.00	0.00	-0.00	-0.00	-0.00
Other	-0.01	-0.00	-0.00	-0.01	-0.01	-0.03	-0.01	-0.00	-0.00	-0.01	-0.01	-0.02	0.00	0.00	0.00	-0.00	-0.00	0.00

5. Comparisons among Items

Effects of Substitution tend to be larger among Items than among subgroups. The Laspeyres-type index became larger than the Paasche-type index in most cases among Items. But some exceptional cases exist. For example, quantity for beef in Fresh meat decreased though the price also decreased from 2000 to 2003, and in the end, the Laspeyres-type index for Fresh meat became lower than the Paasche-type index for the year 2003 (Table 4.1 & 4.2). This case was caused mainly for fear of the BSE (Bovine Spongiform Encephalopathy). Firstly consumers on demand sides departed from the market, and stock farmers or companies on supply sides were forced to lower the price. Similar things happened in Charges for practicing golf. Popularity of playing golf diminished and managers for golf practice lowered the charges to keep customers.

Table 4.1 Laspeyres and Paasche Indices by Yearly Income Quintile Group for Fresh meat for the Year 2003 (2000=100)

	All	I	II	III	IV	V
Laspeyres	101.7	101.8	101.8	101.8	101.7	101.5
Paasche	101.9	102.0	101.9	102.0	101.9	101.6
(P-L)×100/L	0.2	0.2	0.2	0.2	0.2	0.1

Table 4.2 Change of Price and Quantity by Yearly Income Quintile Group for Fresh meat for the Year 2003 (2000=100)

		Change from 2000 to 2003						
	Expenditure in 2000	Price	Quantity					
			All	I	II	III	IV	V
	thousand yen	%	%	%	%	%	%	%
All	63	1.7	-9.3	-7.5	-7.7	-10.5	-10.7	-9.4
Beef	28	-1.1	-17.2	-15.9	-15.3	-19.7	-19.9	-15.1
Pork	23	2.7	-1.7	-0.0	-1.8	-1.4	-2.8	-1.8
Chiken	11	6.5	-5.3	-3.1	-1.7	-7.7	-3.8	-8.4
Others	0	4.0	-15.4	-13.7	-8.9	-18.3	-20.6	-14.7

Then we compared these Laspeyres Indices with geometric mean indices. From natures of the formulae, fixed weights at the base period would value more excessively price down in the geometric formula than in arithmetic formula, if both of the price and the quantity decrease. In this case, the Laspeyres type geometric indices with fixed weight at the base period are less than the Laspeyres type arithmetic indices. Conversely, the Paasche type geometric indices with fixed weights at the reference period are less than the Paasche type arithmetic indices. In the end, Both of geometric indices depart from ideal indices like the Fisher indices or the Tronqvist indices with wider distances than arithmetic indices, what we call in general, the Laspeyres indices and the Paasche indices.

Table 4.3 illustrates the result. Geometric mean indices record lower than arithmetic mean indices if weights are fixed at the base period. On the other hand, geometric mean indices record higher than arithmetic mean indices if weights are fixed at the reference period.

Table 4.3 Comparisons with Indices by Yearly Income Quintile Group for Fresh Meat for the Year 2003 (2000=100)

		All	I	II	III	IV	V
Arithmetic Mean	Laspeyres	101.7	101.8	101.8	101.8	101.7	101.5
	Paasche	101.9	102.0	101.9	102.0	101.9	101.6
	Fisher	101.8	101.9	101.8	101.9	101.8	101.6
	(P-L)×100/L	0.2	0.2	0.2	0.2	0.2	0.1
Geometric Mean	Laspeyres	101.7	101.8	101.7	101.8	101.7	101.5
	Paasche	101.9	102.0	102.0	102.0	102.0	101.7
	Tronqvist	101.8	101.9	101.8	101.9	101.8	101.6
	(P-L)×100/L	0.2	0.3	0.3	0.3	0.3	0.2

Though the Laspeyres-type index became larger than the Paasche-type index for all, reverse results can be seen for some groups. Table 5.1 & 5.2 show that decrease of price and quantity for Fruit or vegetable juice is comparatively large and the Laspeyres-type index is smaller than the Paasche-type index for the subgroup, Other beverages[25] in the group III in the year 2003.

Table 5.1 Laspeyres and Paasche Indices by Yearly Income Quintile Group for Other beverages for the Year 2003 (2000=100)

	All	I	II	III	IV	V
Laspeyres	94.4	94.5	94.5	94.4	94.4	94.4
Paasche	94.4	94.5	94.4	94.5	94.3	94.4
(P-L)×100/L	-0.0	-0.0	-0.1	0.1	-0.1	-0.1

Table 5.2 Change of Price and Quantity by Yearly Income Quintile Group for Other beverages for the Year 2003 (2000=100)

		Change from 2000 to 2003						
	Expenditure in 2000	Price	Quantity					
			All	I	II	III	IV	V

[25] Beverages are composed of Tea, Coffee & cocoa and Other beverages while Alcoholic beverages represents another independent subgroup.

	thousand yen	%	%	%	%	%	%	%
All	**24**	**-5.6**	**4.5**	**4.9**	**6.6**	**2.3**	**2.6**	**6.4**
Fruit or vegetable juice	**14**	**-5.8**	**-2.3**	**2.2**	**-1.0**	**-7.2**	**-0.7**	**-2.9**
Carbonated beverages	**3**	**-8.3**	**8.4**	**-1.6**	**10.6**	**8.5**	**4.4**	**19.7**
Fermented lactic drinks	**4**	**-1.4**	**5.1**	**2.4**	**6.1**	**19.1**	**-7.4**	**6.5**
Others	**3**	**-7.1**	**30.9**	**30.0**	**41.2**	**19.0**	**30.2**	**34.8**

In this case, differences among indices are trivial. But the geometric mean index with fixed weight at the year 2000 records 94.36, slightly smaller than the arithmetic mean index recording 94.38. And difference between the Laspeyres type and the Paasche type is 0.13 numerated by geometric mean indices, slightly larger than 0.08 of that by arithmetic mean indices.

Table 5.3 Coparisons with Iindexes by Yearly Income Quintile Group for Other Beverages for the Year 2003 (2000=100)

□		All	I	II	III	IV	V
Arithmetic Mean	**Laspeyres**	**94.4**	**94.5**	**94.5**	**94.4**	**94.4**	**94.4**
	Paasche	**94.4**	**94.5**	**94.4**	**94.5**	**94.3**	**94.4**
	Fisher	**94.4**	**94.5**	**94.5**	**94.4**	**94.4**	**94.4**
	(P-L)×100/L	**-0.0**	**-0.0**	**-0.1**	**0.1**	**-0.1**	**-0.1**
Geometric Mean	**Laspeyres**	**94.4**	**94.5**	**94.5**	**94.4**	**94.4**	**94.4**
	Paasche	**94.4**	**94.5**	**94.5**	**94.5**	**94.3**	**94.4**
	Tronqvist	**94.4**	**94.5**	**94.5**	**94.4**	**94.4**	**94.4**
	(P-L)×100/L	**0.0**	**0.0**	**-0.0**	**0.1**	**-0.1**	**-0.0**

6. Tentative Conclusion

Above analyses through real CPI in Japan underpin various substitution behaviors such as non-substitution based on institutional factors like medical care and those by economic or social trend sometimes due to specific phenomena like the BSE other than general reversible movement of quantity to prices. The geometric index does not perform well enough as cost-of-living index under uncommon phenomena where substitution rarely take place.

Especially uncommon substitutions tend to happen more in smaller groups like in an income group, because the more similar members within a group are, the more special characteristics as to consumption they tend to have. The situation can be observed in, for instance, expenditures for recreational durables and other beverages for the third income group around the year 2003.

These results suggest that approximation of cost-of-living index, theoretically defined as to keep certain level of utility, from existing data, needs enough care when dealing with prices or expenditures with such characteristics, especially when segmented by household types such as income groups.

The observations however may have been produced under the “moderate deflation”, when subtle movements of both prices and expenditures induce much concerns. Further study would be required to judge whether they are special characteristics only under the present circumstances or will also happen under inflationary trend.

Pursiainen, Heikki: The algebraic interpretation of consistency in aggregation and quasilinear index numbers

See: http://ethesis.helsinki.fi/julkaisut/val/kansa/vk/pursiainen/

Session 3: Housing

Moderator: Keith Woolford

Summary of Session

The measurement of housing prices, in particular those that are relevant to owner-occupiers, is one of the most topical subjects for price statisticians. The requirements for such measures are many and varied and often extend beyond what may be required simply to meet the objectives of official consumer price indexes.

In the case of owner-occupied housing, the policy interest could be in monitoring changes in:

- The full cost of acquiring dwellings;
- The cost of the flow of services derived from the dwellings; and/or
- The out-of-pocket or unrecoverable payments required to gain access to dwellings.

The classes of price indexes relevant to each of these are referred to in turn as the *acquisitions*, *use* and *payments* approaches. The acquisitions approach can be employed to measure changes in house prices either inclusive or exclusive of the land component. There are also two variants of the cost of use approach – the rental equivalence approach and the user cost approach.

There were five papers presented in this session.

The first paper, by Diewert, provides a conceptual treatise of the various options for the measurement of the changes in the cost of the flow of consumption services in each period.

The second paper, by Gudnason, describes the simple user cost methodology used in the Icelandic CPI.

The third paper, by Johannessen, explains why the rental equivalence approach is preferred for the Norwegian CPI.

The fourth paper, by Koev, describes the methodology used to construct the official Finnish House Price Index – using a combination of stratification and hedonic quality adjustment. It also provides an excellent summary of alternative approaches for the construction of house price indexes.

The fifth paper, by Linz and Behrmann, describes the development of an hedonic price index for owner-occupied houses (including land) in Germany.

The papers and the discussion that followed highlighted the conceptual and methodological complexities that need to be addressed to construct reliable price indexes for housing. There is no single correct approach because there is no single correct question. Even with a single agreed objective, measurement practices are likely to vary from country to country to account for differences in housing markets and availability of source data.

Recommendations for Statistical Agencies

1. The production of reliable price measures for owner-occupied housing is not straightforward. Statistical agencies are encouraged to identify the specific requirements of domestic users to assist in the selection of the appropriate measurement objective. There may well be a requirement for more than one measure.

Erwin Diewert:[26] *The Treatment of Owner Occupied Housing and Other Durables in a Consumer Price Index*

Abstract:

The paper systematically surveys alternative approaches to the treatment of durable goods in a consumer price index. The main approaches are the acquisitions, rental equivalence, user cost and payments approaches. A major component of the user cost approach to valuing the services of a durable is the depreciation component. The paper presents a general model of depreciation and then it is specialized to the three most common models of depreciation that are in use. A complication is many durables (like housing) are custom produced and thus the usual methods for determining the form of depreciation are not applicable. The special problems caused by these uniquely produced consumer durables are considered as well as some of the special problems involved in implementing the user cost and rental equivalence methods for valuing the services provided by Owner Occupied Housing.

Key Words:

Durable goods, Consumer Price Index, Cost of Living Index, Owner Occupied Housing, depreciation, hedonic regression models, rental equivalence approach, acquisitions approach, user cost approach, payments approach, maintenance and repair, renovations expenditures.

Journal of Economic Literature Classification Numbers:

C23, C43, C81, D12, E31.

Table of Contents

[26] This research was supported by Statistics Sweden and a Social Sciences and Humanities Research Council of Canada grant. The author thanks Bert Balk, Kevin Fox, Rosmundur Gudnason, Peter Hill, Johannes Hoffman, Arnold Katz, Anders Klevmarken, Timo Koskimäki, Alice Nakamura, Marshall Reinsdorf and Carmit Schwartz for helpful comments on earlier versions of this paper. The above people and institutions are not responsible for any errors or opinions expressed in this paper.

1. Introduction

When a durable good (other than housing) is purchased by a consumer, national Consumer Price Indexes typically attribute *all* of that expenditure to the period of purchase, even though the use of the good extends beyond the period of purchase.[27] This is known as the *acquisitions approach* to the treatment of consumer durables in the context of determining a pricing concept for the CPI. However, if one takes a *cost of living* approach to the Consumer Price Index, then it may be more appropriate to take the cost of *using* the services of the durable good during the period under consideration as the pricing concept. There are two broad methods for estimating this imputed cost for using the services of a durable good during a period:

- If rental or leasing markets for a comparable consumer durable exist, then this market rental price could be used as an estimate for the cost of using the durable during the period. This method is known as the *rental equivalence approach*.

[27] This treatment of the purchases of durable goods dates back to Alfred Marshall (1898; 594-595) at least: "We have noticed also that though the benefits which a man derives from living in his own house are commonly reckoned as part of his real income, and estimated at the net rental value of his house; the same plan is not followed with regard to the benefits which he derives from the use of his furniture and clothes. It is best here to follow the common practice, and not count as part of the national income or dividend anything that is not commonly counted as part of the income of the individual."

- If used or second hand markets for the durable exist, then the imputed cost of purchasing a durable good at the beginning of the period and selling it at the end could be computed and this net cost could be used as a estimate for the cost of using the durable during the period. This method is known as the *user cost approach*.

The major *advantages* of the acquisitions approach to the treatment of consumer durables are:

- It is conceptually simple and entirely similar to the treatment of nondurables and services and
- No complex imputations are required.

The major *disadvantage* of the acquisitions approach compared to the other two approaches is that the acquisitions approach is not likely to reflect accurately the consumption services of consumer durables in any period. Thus suppose that real interest rates in a country become very high due to some sort of macroeconomic crisis. Under these conditions, typically purchases of automobiles and houses and other long lived consumer durables drop dramatically, perhaps to zero. However, the actual consumption of automobile and housing services of the country's population will not fall to zero under these circumstances: consumers will still be consuming the services of their existing stocks of autos and houses. Thus for at least some purposes, rather than taking the cost of *purchasing* a consumer durable as the pricing concept, it will be more useful to take the cost of *using* the services of the durable good during the period under consideration as the pricing concept.

The above paragraphs provide a brief overview of the three major approaches to the treatment of consumer durables. In the remainder of this introduction, we explore these approaches in a bit more detail and give the reader an outline of the detailed discussion that will follow in subsequent sections.

We first consider a formal definition of a consumer durable. By definition, a durable good delivers services longer than the period under consideration.[28] The *System of National Accounts 1993* defines a *durable good* as follows:

"In the case of goods, the distinction between acquisition and use is analytically important. It underlies the distinction between durable and non-durable goods extensively used in economic analysis. In fact, the distinction between durable and non-durable goods is not based on physical durability as such. Instead, the distinction is based on whether the goods can be used once only for purposes of production or consumption or whether they can be used repeatedly, or continuously. For example, coal is a highly durable good in a physical sense, but it can be burnt only once. A durable good is therefore defined as one which may be used repeatedly or continuously over a period of more than a year, assuming a normal or average rate of physical usage. A consumer durable is a good that may be used for purposes of consumption repeatedly or continuously over a period of a year or more." *System of National Accounts 1993*, (1993; 208).

According to the above National Accounts definition, durability is more than the fact that a good can physically persist for more than a year (this is true of most goods): a durable good is distinguished from a nondurable good due to its property that it can deliver useful services to a consumer through repeated use over an extended period of time.

Since the benefits of using the consumer durable extend over more than one period, it does not seem to be appropriate to charge the entire purchase cost of the durable to the initial period of purchase.

[28] An alternative definition of a durable good is that the good delivers services to its purchaser for a period exceeding three years: "The Bureau of Economic Analysis defines consumer durables as those durables that have an average life of at least 3 years." Arnold J. Katz (1983; 422).

If this point of view is taken, then the initial purchase cost must be distributed somehow over the useful life of the asset. This is *a fundamental problem of accounting*.[29] Hulten (1990) explains the consequences for accountants of the durability of a purchase as follows:

"Durability means that a capital good is productive for two or more time periods, and this, in turn, implies that a distinction must be made between the value of using or renting capital in any year and the value of owning the capital asset. This distinction would not necessarily lead to a measurement problem if the capital services used in any given year were paid for in that year; that is, if all capital were rented. In this case, transactions in the rental market would fix the price and quantity of capital in each time period, much as data on the price and quantity of labor services are derived from labor market transactions. But, unfortunately, much capital is utilized by its owner and the transfer of capital services between owner and user results in an implicit rent typically not observed by the statistician. Market data are thus inadequate for the task of directly estimating the price and quantity of capital services, and this has led to the development of indirect procedures for inferring the quantity of capital, like the perpetual inventory method, or to the acceptance of flawed measures, like book value." Charles R. Hulten (1990; 120-121).

Thus the treatment of durable goods is more complicated than the treatment of nondurable goods and services due to the simple fact that the period of time that a durable is used by the consumer extends beyond the period of purchase. For nondurables and services, the price statistician's measurement problems are conceptually simple: prices for the same commodity need only be collected in each period and compared. However, for a durable good, the periods of payment and use do not coincide and so complex imputation problems arise if the goal of the price statistician is to measure and compare the price of *using* the services of the durable in two time periods.

As mentioned above, there are 3 main methods for dealing with the durability problem:

- Ignore the problem of distributing the initial cost of the durable over the useful life of the good and allocate the entire charge to the period of purchase. As noted above, this is known as the *acquisitions approach* and it is the present approach used by Consumer Price Index statisticians for all durables with the exception of housing.
- The *rental equivalence approach*. In this approach, a period price is imputed for the durable which is equal to the rental price or leasing price of an equivalent consumer durable for the same period of time.
- The *user cost approach.* In this approach, the initial purchase cost of the durable is decomposed into two parts: one part which reflects an estimated cost of using the services of the durable for the period and another part, which is regarded as an investment, which must earn some exogenous rate of return.

These three major approaches will be discussed more fully in sections 2, 3 and 4 below. However, there is a fourth approach to the treatment of consumer durables that has only been used in the context of pricing owner occupied housing and that is the *payments approach* [30]. This is a kind of a cash flow approach, which is not entirely satisfactory. It will be briefly discussed in section 12 after we have discussed the treatment of owner occupied housing in more detail.

[29] "The third convention is that of the annual accounting period. It is this convention which is responsible for most of the difficult accounting problems. Without this convention, accounting would be a simple matter of recording completed and fully realized transactions: an act of primitive simplicity." Stephen Gilman (1939; 26).
"*All* the problems of income measurement are the result of our desire to attribute income to arbitrarily determined short periods of time. Everything comes right in the end; but by then it is too late to matter." David Solomons (1961; 378). Note that these authors do not mention the additional complications that are due to the fact that future revenues and costs must be discounted to yield values that are equivalent to present dollars.
[30] This is the term used by Goodhart (2001; F350-F351).

The above three approaches to the treatment of durable purchases can be applied to the purchase of any durable commodity. However, historically, it turns out that the rental equivalence and user cost approaches have *only* been applied to owner occupied housing. In other words, the acquisitions approach to the purchase of consumer durables has been universally used by statistical agencies, with the exception of owner occupied housing. A possible reason for this is tradition; i.e., Marshall set the standard and statisticians have followed his example for the past century. However, another possible reason is that unless the durable good has a very long useful life, it usually will not make a great deal of difference in the long run whether the acquisitions approach or one of the two alternative approaches is used. This point is discussed in more detail in section 5 below.

A major component of the user cost approach to valuing the services of owner occupied housing is the depreciation component. In section 6, a general model of depreciation for a consumer durable is presented and then it is specialized to the three most common models of depreciation that are in use. The models presented in section 6 assume that homogeneous units of the durable are produced in each period so that information on the prices of the various vintages of the durable at any point in time can be used to determine the pattern of depreciation. However, many durables (like housing) are custom produced and thus the methods for determining the form of depreciation explained in section 6 are not applicable. The special problems caused by these uniquely produced consumer durables are considered in section 7.

Sections 8, 9, 10 and 11 treat some of the special problems involved in implementing the user cost and rental equivalence methods for valuing the services provided by Owner Occupied Housing (OOH). Section 8 presents a derivation for the user cost of OOH and various approximations to it. Section 9 looks at some of the problems associated with obtaining constant quality prices for housing. Section 10 considers some of the costs that are tied to home ownership while section 11 considers how a landlord's costs might differ from a homeowner's costs. This material is relevant if the rental equivalence approach to valuing the services of OOH is used: care must be taken to remove some costs that are imbedded in market rents that homeowners do not face.

Section 13 tries to bring together all of the material on the problems associated with pricing Owner Occupied Housing and to outline possible CPI measurement strategies.

2. The Acquisitions Approach

The *net acquisitions approach* to the treatment of owner occupied housing is described by Goodhart as follows:

"The first is the net acquisition approach, which is the change in the price of newly purchased owner occupied dwellings, weighted by the net purchases of the reference population. This is an asset based measure, and therefore comes close to my preferred measure of inflation as a change in the value of money, though the change in the price of the stock of existing houses rather than just of net purchases would in some respects be even better. It is, moreover, consistent with the treatment of other durables. A few countries, e.g., Australia and New Zealand, have used it, and it is, I understand, the main contender for use in the Euro-area Harmonized Index of Consumer Prices (HICP), which currently excludes any measure of the purchase price of (new) housing, though it does include minor repairs and maintenance by home owners, as well as all expenditures by tenants." Charles Goodhart (2001; F350).

Thus the weights for the net acquisitions approach are the net purchases of the household sector of houses from other institutional sectors in the base period. Note that in principle, purchases of second-hand dwellings from other sectors are relevant here; e.g., a local government may sell rental dwellings to owner occupiers. However, typically, newly built houses form a major part of these types of transactions. Thus the long term price relative for this category of expenditure will be primarily the price of (new) houses (quality adjusted) in the current period relative to the price of new houses in the base period.[31] If this approach is applied to other consumer durables, it is extremely easy to implement: the purchase of a durable is treated in the same way as a nondurable or service purchase is treated.

One additional implication of the net acquisition approach is that major renovations and additions to owner occupied dwelling units could also be considered as being in scope for this approach. In practice, these costs typically are not covered in a standard consumer price index. The treatment of renovations and additions will be considered in more detail in section 10.4 below.

Traditionally, the net acquisitions approach also includes transfer costs relating to the buying and selling of second hand houses as expenditures that are in scope for an acquisitions type consumer price index. These costs are mainly the costs of using a real estate agent's services and asset transfer taxes. These transfer costs will be further discussed in sections 10.2 and 10.5 below.

The major advantage of the acquisitions approach is that it treats durable and nondurable purchases in a completely symmetric manner and thus no special procedures have to be developed by a statistical agency to deal with durable goods. As will be seen in section 5 below, the major disadvantage of this approach is that the *expenditures* associated with this approach will tend to *understate* the corresponding expenditures on durables that are implied by the rental equivalence and user cost approaches.

Some differences between the acquisitions approach and the other approaches are:

- If rental or leasing markets for the durable exist and the durable has a long useful life, then the expenditure weights implied by the rental equivalence or user cost approaches will typically be much larger than the corresponding expenditure weights implied by the acquisitions approach; see Section 5 below.
- If the base year corresponds to a boom year (or a slump year) for the durable, then the base period expenditure weights may be too large or too small. Put another way, the aggregate expenditures that correspond to the acquisitions approach are likely to be more volatile than the expenditures for the aggregate that are implied by the rental equivalence or user cost approaches.
- In making comparisons of consumption across countries where the proportion of owning versus renting or leasing the durable varies greatly,[32] the use of the acquisitions approach may lead to misleading cross country comparisons. The reason for this is that opportunity costs of capital are excluded in the net acquisitions approach whereas they are explicitly or implicitly included in the other two approaches.

More fundamentally, whether the acquisitions approach is the right one or not depends on the overall purpose of the index number. If the purpose is to measure the price of current period *consumption services*, then the acquisitions approach can only be regarded as an approximation to a more appropriate approach (which would be either the rental equivalence or user cost approach). If

[31] This price index may or may not include the price of the land that the new dwelling unit sits on; e.g., a new house price construction index would typically not include the land cost. The acquisitions approach concentrates on the purchases by households of goods and services that are provided by suppliers from outside the household sector. Thus if the land on which a new house sits was previously owned by the household sector, then presumably, the cost of this land would be excluded from an acquisitions type new house price index.

[32] From Hoffmann and Kurz (2002; 3-4), about 60% of German households live in rented dwellings whereas only about 11% of Spaniards rent their dwellings in 1999 (private communication).

the purpose of the index is to measure *monetary* (or nonimputed) *expenditures* by households during the period, then the acquisitions approach is preferable, since the rental equivalence and user cost approaches necessarily involve imputations.

3. The Rental Equivalence Approach

The *rental equivalence approach* simply values the services yielded by the use of a consumer durable good for a period by the corresponding market rental value for the same durable for the same period of time (if such a rental value exists). This is the approach taken in the *System of National Accounts: 1993* for owner occupied housing:

"As well-organized markets for rented housing exist in most countries, the output of own-account housing services can be valued using the prices of the same kinds of services sold on the market with the general valuation rules adopted for goods and services produced on own account. In other words, the output of housing services produced by owner-occupiers is valued at the estimated rental that a tenant would pay for the same accommodation, taking into account factors such as location, neighbourhood amenities, etc. as well as the size and quality of the dwelling itself." Eurostat and others (1993; 134).

However, the *System of National Accounts: 1993* follows Marshall (1898; 595) and does *not* extend the rental equivalence approach to consumer durables other than housing. This seemingly inconsistent treatment of durables is explained in the *SNA 1993* as follows:

"The production of housing services for their own final consumption by owner-occupiers has always been included within the production boundary in national accounts, although it constitutes an exception to the general exclusion of own-account service production. The ratio of owner-occupied to rented dwellings can vary significantly between countries and even over short periods of time within a single country, so that both international and intertemporal comparisons of the production and consumption of housing services could be distorted if no imputation were made for the value of own-account services." Eurostat and others (1993; 126).

Eurostat's (2001) *Handbook on Price and Volume Measures in National Accounts* also recommends the rental equivalence approach for the treatment of the dwelling services for owner occupied housing:

"The output of dwelling services of owner occupiers at current prices is in many countries estimated by linking the actual rents paid by those renting similar properties in the rented sector to those of owner occupiers. This allows the imputation of a notional rent for the service owner occupiers receive from their property." Eurostat (2001; 99).

The US statistical agencies, the Bureau of Labor Statistics and the Bureau of Economic Analysis, both use the rental equivalence approach to value the services of owner occupied housing. Katz describes the BEA procedures as follows:

"Basically, BEA measures the gross rent (space rent) of owner occupied housing from data on the rent paid for similar housing with the same market value. To get the service value that is added to GNP (gross housing product), the value of intermediate goods and services included in this figure (e. g., expenditures for repair and maintenance, insurance, condominium fees, and closing costs) are subtracted from the space rent. To obtain a net return (net rental income), depreciation, taxes, and net interest are subtracted from, and subsidies added to, the service value." Arnold J. Katz (1983; 411).

There are some problems with the above treatment of housing and they will be discussed in later sections after the user cost approach to durables has been discussed.[33]

To summarize the above material, it can be seen that the rental equivalence approach to the treatment of durables is conceptually simple: impute a current period rental or leasing price for a comparable product as the price for the purchase of a unit of a consumer durable. For existing stocks of used consumer durables, the rental equivalence approach would entail finding rental prices for comparable used units.[34] To date, as noted above, statistical agencies have not done this, with the single exception of owner occupied housing. However, note that in order to implement the rental equivalence approach, it is necessary that the relevant rental or leasing markets exist and often this will not be the case, particularly when it is recognized that vintage specific rental prices may be required for all vintages of the durable held by households.[35]

4. The User Cost Approach

The user cost approach to the treatment of durable goods is in some ways very simple: it calculates the cost of purchasing the durable at the beginning of the period, using the services of the durable during the period and then netting off from these costs the benefit that could be obtained by selling the durable at the end of the period. However, there are several details of this procedure that are somewhat controversial. These details involve the use of opportunity costs, which are usually imputed costs, the treatment of interest and the treatment of capital gains or holding gains.

Another complication with the user cost approach is that it involves making distinctions between current period (flow) purchases within the period under consideration and the holdings of physical stocks of the durable at the beginning and the end of the accounting period. Up to this point, all prices and quantity purchases were thought of as taking place at a single point in time, say the middle of the period under consideration, and consumption was thought of as taking place within the period as well. Thus, there was no need to consider the behavior (and valuation) of stocks of consumer durables that households may have a their disposal. The rather complex problems involved in accounting for stocks and flows are unfamiliar to most price statisticians.

To determine the net cost of using the durable good during say period 0, assume that one unit of the durable good is purchased at the beginning of period 0 at the price P^0. The "used" or "second-hand" durable good can be sold at the end of period 0 at the price P_S^1. It might seem that a reasonable net cost for the use of one unit of the consumer durable during period 0 is its initial purchase price P^0 less its end of period 0 "scrap value" P_S^1. However, *money received at the end of the period is not as valuable as money that is received at the beginning of the period*. Thus in order to convert the end of period value into its beginning of the period equivalent value, it is necessary to *discount* the

[33] To anticipate the later results: the main problem is that the rental equivalence approach to valuing the services of owner occupied housing may give a higher valuation for these services than the user cost approach.

[34] Another method for determining rental price equivalents for stocks of consumer durables is to *ask* households what they think their durables would rent for. This approach is used by the Bureau of Labor Statistics in order to determine expenditure weights for owner occupied housing; i.e., homeowners are asked to estimate what their house would rent for if it were rented to a third party; see the Bureau of Labor Statistics (1983). Lebow and Rudd (2003; 169) note that these consumer expenditure survey based estimates of imputed rents in the US differ considerably from the corresponding Bureau of Economic Analysis estimates for imputed rents, which are based on applying a rent to value ratio for rented properties to the owner occupied stock of housing. Lebow and Rudd feel that the expenditure survey estimates may be less reliable than ratio of rent to value method due to the relatively small size of the consumer expenditure survey plus the difficulties households may have in recalling or estimating expenditures.

[35] However, if the form of depreciation is of the one hoss shay or light bulb type, then the rental price for the durable will be the same for all vintages and hence a detailed knowledge of market rentals by vintage will not be required. The light bulb model of depreciation dates back to Böhm-Bawerk (1891; 342). For more recent material on this model, see section 6.4 below or Hulten (1990) or Diewert (2003b).

term P_S^1 by the term $1+r^0$ where r^0 is the beginning of period 0 nominal interest rate that the consumer faces. Hence the *period 0 user cost* u^0 for the consumer durable[36] is defined as

$$(1)\ u^0 \equiv P^0 - P_S^1/(1+r^0)\ .$$

There is another way to view the user cost formula (1): the consumer purchases the durable at the beginning of period 0 at the price P^0 and charges himself or herself the rental price u^0. The remainder of the purchase price, I^0, defined as

$$(2)\ I^0 \equiv P^0 - u^0$$

can be regarded as an *investment*, which is to yield the appropriate opportunity cost of capital r^0 that the consumer faces. At the end of period 0, this rate of return could be realized provided that I^0, r^0 and the selling price of the durable at the end of the period P_S^1 satisfy the following equation:

$$(3)\ I^0(1+r^0) = P_S^1\ .$$

Given P_S^1 and r^0, (3) determines I^0, which in turn, given P^0, determines the user cost u^0 via (2)[37].

It should be noted that some price statisticians object to the user cost concept as a valid pricing concept for a Consumer Price Index:

"A suitable price concept for a CPI ought to reflect only a ratio of exchange of money for other things, not a ratio at which money in one form or time period can be traded for money in another form or time period. The ratio at which money today can be traded for money tomorrow by paying an interest rate or by enjoying actual or expected holding gains on an appreciating asset has no part in a measure of the current purchasing power of money." Marshall Reinsdorf (2003).

Thus user costs are not like the prices of nondurables or services because the user cost concept involves pricing the durable at *two* points in time rather than at a single point in time.[38] Because the user cost concept involves prices at two points in time, money received or paid out at the first point in time is more valuable than money paid out or received at the second point in time and so *interest rates* creep into the user cost formula. Furthermore, because the user cost concept involves prices at two points in time, *expected prices* can be involved if the user cost is calculated at the beginning of the period under consideration instead of at the end. With all of these complications, it is no wonder that many price statisticians would like to avoid the using user costs as a pricing concept. However, even for price statisticians who would prefer to use the rental equivalence approach to the treatment of durables over the user cost approach, there is some justification for considering the user cost approach in some detail, since this approach gives insights into the economic determinants of the rental or leasing price of a durable. As will be seen in section 11 below, the user cost for a

---

[36] This approach to the derivation of a user cost formula was used by Diewert (1974b) who in turn based it on an approach due to Hicks (1946; 326).

[37] This derivation for the user cost of a consumer durable was also made by Diewert (1974b; 504).

[38] Woolford also suggested that interest should be excluded from an ideal price index that measured inflation. In his view, interest is not a *contemporaneous price*; i.e., an interest rate necessarily refers to *two* points in time; a beginning point when the capital is loaned and an ending point when the capital loaned must be repaid. Thus if one wanted to restrict attention to a domain of definition that consisted of only contemporaneous prices, interest rates would be excluded. Woolford (1999; 535) noted that his ideal inflation measure "would be contemporary in nature, capturing only the current trend in prices associated with transactions in goods and services. It would exclude interest rates on the ground that they are intertemporal prices, representing the relative price of consuming today rather than in the future."

house can differ substantially for a landlord compared to an owner and thus adjustments should be made to market rents for dwelling units if these observed rents are to be used as imputations for owner occupied rents.

The user cost formula (1) can be put into a more familiar form if the period 0 *economic depreciation rate* δ and the period 0 *ex post asset inflation rate* i^0 are defined. Define δ by:

(4) $(1 - \delta) \equiv P_S^1/P^1$

where P_S^1 is the price of a used asset at the end of period 0 and P^1 is the price of a new asset at the end of period 0. The *period 0 inflation rate* for the new asset, i^0, is defined by:

(5) $1+i^0 \equiv P^1/P^0$.

Eliminating P^1 from equations (4) and (5) leads to the following formula for the end of period 0 used asset price:

(6) $P_S^1 = (1 - \delta)(1 + i^0)P^0$.

Substitution of (6) into (1) yields the following expression for the *period 0 user cost* u^0:

(7) $u^0 = [(1 + r^0) - (1 - \delta)(1 + i^0)]P^0 / (1 + r^0)$.

Note that $r^0 - i^0$ can be interpreted as a period 0 *real interest rate* and $\delta(1+i^0)$ can be interpreted as an *inflation adjusted depreciation rate*.

The user cost u^0 is expressed in terms of prices that are discounted to the *beginning* of period 0. However, it is also possible to express the user cost in terms of prices that are "discounted" to the *end* of period 0.[39] Thus define the *end of period 0 user cost* p^0 as:[40]

(8) $p^0 \equiv (1 + r^0)u^0 = [(1 + r^0) - (1 - \delta)(1 + i^0)]P^0$

where the last equation follows using (7). If the real interest rate r^{0*} is defined as the nominal interest rate r^0 less the asset inflation rate i^0 and the small term δi^0 is neglected, then the end of the period user cost defined by (8) reduces to:

(9) $p^0 = (r^{0*} + \delta)P^0$.

[39] Thus the beginning of the period user cost u^0 discounts all monetary costs and benefits into their dollar equivalent at the beginning of period 0 whereas p^0 discounts (or appreciates) all monetary costs and benefits into their dollar equivalent at the end of period 0. This leaves open how flow transactions that take place within the period should be treated. Following the conventions used in financial accounting suggests that flow transactions taking place within the accounting period be regarded as taking place at the end of the accounting period and hence following this convention, end of period user costs should be used by the price statistician.

[40] Christensen and Jorgenson (1969) derived a user cost formula similar to (7) in a different way using a continuous time optimization model. If the inflation rate i equals 0, then the user cost formula (7) reduces to that derived by Walras (1954; 269) (first edition 1874). This zero inflation rate user cost formula was also derived by the industrial engineer A. Hamilton Church (1901; 907-908), who perhaps drew on the work of Matheson: "In the case of a factory where the occupancy is assured for a term of years, and the rent is a first charge on profits, the rate of interest, to be an appropriate rate, should, so far as it applies to the buildings, be equal (including the depreciation rate) to the rental which a landlord who owned but did not occupy a factory would let it for." Ewing Matheson (1910; 169), first published in 1884. Additional derivations of user cost formulae in discrete time have been made by Katz (1983; 408-409) and Diewert (2003b). Hall and Jorgenson (1967) introduced tax considerations into user cost formulae.

Abstracting from transactions costs and inflation, it can be seen that the end of the period user cost defined by (9) is an *approximate rental cost*; i.e., the rental cost for the use of a consumer (or producer) durable good should equal the (real) opportunity cost of the capital tied up, r^0*P^0, plus the decline in value of the asset over the period, δP^0. Formulae (8) and (9) thus cast some light on what are the economic determinants of rental or leasing prices for consumer durables.

If the simplified user cost formula defined by (9) above is used, then forming a price index for the user costs of a durable good is not very much more difficult than forming a price index for the purchase price of the durable good, P^0. The price statistician needs only to:

- Make a reasonable assumption as to what an appropriate monthly or quarterly real interest rate r^0* should be;
- Make an assumption as to what a reasonable monthly or quarterly depreciation rate δ should be;[41]
- Collect purchase prices P^0 for the durable and
- Make an estimate of the total stock of the durable which was held by the reference population during the base period for quantities. In order to construct a superlative index, estimates of the stock held will have to be made for each period.

If it is thought necessary to implement the more complicated user cost formula (8) in place of the simpler formula (9), then the situation is more complicated. As it stands, the end of the period user cost formula (8) is an *ex post* (or after the fact) *user cost*: the asset inflation rate i^0 cannot be calculated until the end of period 0 has been reached. Formula (8) can be converted into an *ex ante* (or before the fact) *user cost* formula if i^0 is interpreted as an *anticipated asset inflation rate*. The resulting formula should approximate a market rental rate for the asset under inflationary conditions.[42]

Note that in the user cost approach to the treatment of consumer durables, the *entire* user cost formula (8) or (9) is the period 0 price. Thus in the time series context, it is *not* necessary to deflate each component of the formula *separately*; the period 0 price $p^0 \equiv [r^0 - i^0 + \delta(1+i^0)]P^0$ is compared to the corresponding period 1 price, $p^1 \equiv [r^1 - i^1 + \delta(1+i^1)]P^1$ and so on.

In principle, depreciation rates can be estimated using information on the selling prices of used units of the durable good. In section 6 below, this methodology will be explained in more detail. However, before this is done, it will be useful to use the material in this section to explain what the relationship between the user cost and acquisition approaches to the treatment of durables is likely to be. This topic is discussed in the following section.

5. The Relationship Between User Costs and Acquisition Costs

In this section, the user cost approach to the treatment of consumer durables will be compared to the acquisitions approach. Obviously, in the short run, the value flows associated with each approach

[41] The geometric model for depreciation to be explained in more detail in section 6.2 below requires only a single monthly or quarterly depreciation rate. Other models of depreciation may require the estimation of a sequence of vintage depreciation rates. If the estimated annual geometric depreciation rate is δ_a, then the corresponding monthly geometric depreciation rate δ can be obtained by solving the equation $(1 - \delta)^{12} = 1 - \delta_a$. Similarly, if the estimated annual real interest rate is r_a*, then the corresponding monthly real interest rate $r*$ can be obtained by solving the equation $(1 + r*)^{12} = 1 + r_a*$.

[42] Since landlords must set their rent at the beginning of the period (and in fact, they usually set their rent for an extended period of time), if the user cost approach is used to model the economic determinants of market rental rates, then the asset inflation rate i^0 should be interpreted as an expected inflation rate rather than an after the fact actual inflation rate. This use of ex ante prices in this price measurement context should be contrasted with the preference of national accountants to use actual or ex post prices in the system of national accounts.

could be very different. For example, if real interest rates, $r^0 - i^0$, are very high and the economy is in a severe recession or depression, then purchases of new consumer durables, Q^0 say, could be very low and even approach 0 for very long lived assets, like houses. On the other hand, using the user cost approach, existing stocks of consumer durables would be carried over from previous periods and priced out at the appropriate user costs and the resulting consumption value flow could be quite large. Thus in the short run, the monetary values of consumption under the two approaches could be vastly different. Hence, in what follows, a (hypothetical) longer run comparison is considered where real interest rates are held constant.[43]

Suppose that in period 0, the reference population of households purchased q^0 units of a consumer durable at the purchase price P^0. Then *the period 0 value of consumption from the viewpoint of the acquisitions approach* is:

(10) $V_A{}^0 \equiv P^0q^0$.

Recall that the end of period user cost for one new unit of the asset purchased at the beginning of period 0 was p^0 defined by (8) above. In order to simplify the analysis, *declining balance depreciation* is assumed[44]; i.e., at the beginning of period 0, a one period old asset is worth $(1-\delta)P^0$; a two period old asset is worth $(1-\delta)^2P^0$; … ; a t period old asset is worth $(1-\delta)^tP^0$; etc. Under these hypotheses, the corresponding end of period 0 user cost for a new asset purchased at the beginning of period 0 is p^0; the end of period 0 user cost for a one period old asset at the beginning of period 0 is $(1-\delta)p^0$; the corresponding user cost for a two period old asset at the beginning of period 0 is $(1-\delta)^2p^0$; … ; the corresponding user cost for a t period old asset at the beginning of period 0 is $(1-\delta)^tp^0$; etc.[45] The final simplifying assumption is that household purchases of the consumer durable have been growing at the geometric rate g into the indefinite past. This means that if household purchases of the durable were q^0 in period 0, then in the previous period they purchased $q^0/(1+g)$ new units; two periods ago, they purchased $q^0/(1+g)^2$ new units; … ; t periods ago, they purchased $q^0/(1+g)^t$ new units; etc. Putting all of these assumptions together, it can be seen that *the period 0 value of consumption from the viewpoint of the user cost approach* is:

(11) $V_U{}^0 \equiv p^0q^0 + [(1-\delta)p^0q^0/(1+g)] + [(1-\delta)^2 p^0q^0/(1+g)^2] + \ldots$
$= (1+g)p^0q^0/(g+\delta)$ summing the infinite series
(12) $= (1+g)[(1+r^0) - (1-\delta)(1+i^0)]P^0q^0/(g+\delta)$ using (8).

Equation (12) can be simplified by letting the asset inflation rate i^0 be 0 (or by replacing $r^0 - i^0$ by the real interest rate r^{0*} and by ignoring the small term δi^0) and under these conditions, the ratio of the user cost flow of consumption (12) to the acquisitions measure of consumption in period 0, (10) is:

(13) $V_U{}^0/V_A{}^0 = (1+g)(r^{0*} + \delta)/(g+\delta)$.

Using formula (13), it can be seen that if $1+g > 0$ and $\delta + g > 0$, then $V_U{}^0 / V_A{}^0$ will be greater than unity if

(14) $r^{0*} > g(1-\delta)/(1+g)$,

[43] The following material is based on Diewert (2002c).
[44] This form of depreciation will be discussed in more detail in section 6.2 below.
[45] For some consumer durables, the one hoss shay assumption for depreciation may be more realistic than the declining balance model; see section 6.4 below or Hulten (1990) or Diewert and Lawrence (2000).

a condition that will usually be satisfied.[46] Thus under normal conditions and over a longer time horizon, *household expenditures on consumer durables using the user cost approach will tend to exceed the corresponding money outlays on new purchases of the consumer durable*. The difference between the two approaches will tend to grow as the life of the asset increases (i.e., as the depreciation rate δ decreases).

To get a rough idea of the possible magnitude of the value ratio for the two approaches, V_U^0/V_A^0, equation (13) is evaluated for a "housing" example using annual data where the depreciation rate is 2 % (i.e., $\delta = .02$), the real interest rate is 4 % (i.e., $r^{0*} = .04$) and the growth rate for the production of new houses is 1 % (i.e., $g = .01$). In this base case, the ratio of user cost expenditures on housing to the purchases of new housing in the same period, V_U^0/V_A^0, is 2.02. If the depreciation rate is *increased* to 3 %, then V_U^0/V_A^0 *decreases* to 1.77; if the depreciation rate is *decreased* to 1 %, then V_U^0/V_A^0 *increases* to 2.53. Again looking at the base case, if the real interest rate is *increased* to 5 %, then V_U^0/V_A^0 *increases* to 2.36 while if the real interest rate is *decreased* to 3 %, then V_U^0/V_A^0 *decreases* to 1.68. Finally, if the growth rate for new houses is *increased* to 2 %, then V_U^0/V_A^0 *decreases* to 1.53 while if the growth rate is *decreased* to 0, then V_U^0/V_A^0 *increases* to 3.00. *Thus an acquisitions approach to housing in the CPI is likely to give about one half the expenditure weight that a user cost approach would give*.

For shorter lived assets, the difference between the acquisitions approach and the user cost approach will not be so large and hence, this justifies the acquisitions approach as being approximately "correct" as a measure of consumption services.[47]

Here is a list of some of the problems and difficulties that might arise in implementing a user cost approach to purchases of a consumer durable:[48]

- It is difficult to determine what the relevant nominal interest rate r^0 is for each household. If a consumer has to borrow to finance the cost of a durable good purchase, then this interest rate will typically be much higher than the safe rate of return that would be the appropriate opportunity cost rate of return for a consumer who had no need to borrow funds to finance the purchase.[49] It may be necessary to simply use a benchmark interest rate that would be determined by either the government, a national statistical agency or an accounting standards board.

[46] Note that if the real interest rate r^0 equals g, the real rate of growth in the purchases of the durable, then from (13), $V_U^0 / V_A^0 = (1+g)$ and the acquisitions approach will be more or less equivalent to the user cost approach over the long run.

[47] The simplified user cost approach can be used for other consumer durables as well. In formula (13), let $r^{0*} = .04$, $g = .01$ and $\delta = .15$ and under these conditions, $V_U^0/V_A^0 = 1.20$; i.e., for a declining balance depreciation rate of 15%, the user cost approach leads to an estimated value of consumption that is 20% higher than the acquisitions approach under the conditions specified. Thus for consumer durable depreciation rates that are lower than 15%, it would be useful for the statistical agency to produce user costs for these goods and for the national accounts division to produce the corresponding consumption flows as "analytic series". It should be noted that this extends the present national accounts treatment of housing to other long lived consumer durables. Note also that this revised treatment of consumption in the national accounts would tend to make rich countries richer, since poorer countries hold fewer long lived consumer durables on a per capita basis.

[48] For additional material on difficulties with the user cost approach, see Diewert (1980; 475-479) and Katz (1983; 415-422).

[49] Katz (1983; 415-416) comments on the difficulties involved in determining the appropriate rate of interest to use: "There are numerous alternatives: a rate on financial borrowings, on savings, and a weighted average of the two; a rate on nonfinancial investments. e.g., residential housing, perhaps adjusted for capital gains; and the consumer's subjective rate of time preference. Furthermore, there is some controversy about whether it should be the maximum observed rate, the average observed rate, or the rate of return earned on investments that have the same degree of risk and liquidity as the durables whose services are being valued."

- It will generally be difficult to determine what the relevant depreciation rate is for the consumer durable.[50]
- *Ex post user costs* based on formula (8) will be too volatile to be acceptable to users[51] (due to the volatility of the asset inflation rate i^0) and hence an *ex ante user cost* concept will have to be used. This creates difficulties in that different national statistical agencies will generally make different assumptions and use different methods in order to construct forecasted structures and land inflation rates and hence the resulting ex ante user costs of the durable may not be comparable across countries.[52]
- The user cost formula (8) should be generalized to accommodate various taxes that may be associated with the purchase of a durable or with the continuing use of the durable.[53]

Some of the problems associated with estimating depreciation rates will be discussed in the next section.

6. Alternative Models of Depreciation

6.1 A General Model of Depreciation for (Unchanging) Consumer Durables

In this subsection, a "general" model of depreciation for durable goods that appear on the market each period without undergoing quality change will be presented. In three subsequent subsections, this general model will be specialized to the three most common models of depreciation that appear in the literature. In section 7 below, the additional problems that occur when the durable is built as a unique good will be discussed.

The main tool that can be used to identify depreciation rates for a durable good is the (cross sectional) sequence of vintage asset prices that units of the good sell for on the second hand market at any point of time.[54]

Some notation is required. Let P^0 be the price of a newly produced unit of the durable good at the beginning of period 0 (this is the same notation as was used earlier). Let P_v^t be the second hand

[50] It is not necessary to assume declining balance depreciation in the user cost approach: any pattern of depreciation can be accommodated, including one hoss shay depreciation, where the durable yields a constant stream of services over time until it is scrapped. See Diewert and Lawrence (2000) for some empirical examples for Canada using different assumptions about the form of depreciation. For references to the depreciation literature and for empirical methods for estimating depreciation rates, see Hulten and Wykoff (1981a) (1981b) (1996) and Jorgenson (1996).

[51] Goodhart (2001; F351) comments on the practical difficulties of using ex post user costs for housing as follows: "An even more theoretical user cost approach is to measure the cost foregone by living in an owner occupied property as compared with selling it at the beginning of the period and repurchasing it at the end ... But this gives the absurd result that as house prices rise, so the opportunity cost falls; indeed the more virulent the inflation of housing asset prices, the more negative would this measure become. Although it has some academic aficionados, this flies in the face of common sense; I am glad to say that no country has adopted this method." As will be seen later, Iceland has in fact adopted a simplified user cost framework.

[52] For additional material on the difficulties involved in constructing ex ante user costs, see Diewert (1980; 475-486) and Katz (1983; 419-420). For empirical comparisons of different user cost formulae, see Harper, Berndt and Wood (1989) and Diewert and Lawrence (2000).

[53] For example, property taxes are associated with the use of housing services and hence should be included in the user cost formula; see section 10.2 below. As Katz (1983; 418) noted, taxation issues also impact the choice of the interest rate: "Should the rate of return be a before or after tax rate?" From the viewpoint of a household that is not borrowing to finance the purchase of the durable, an after tax rate of return seems appropriate but from the point of a leasing firm, a before tax rate of return seems appropriate. This difference helps to explain why rental equivalence prices for the durable might be higher than user cost prices; see also section 11.4 below.

[54] Another information source that could be used to identify depreciation rates for the durable good is the sequence of vintage rental or leasing prices that might exist for some consumer durables. In the closely related capital measurement literature, the general framework for an internally consistent treatment of capital services and capital stocks in a set of vintage accounts was set out by Jorgenson (1989) and Hulten (1990; 127-129) (1996; 152-160).

market price at the beginning of period t of a unit of the durable good that is v periods old.[55] Let δ_v^0 be the period 0 depreciation rate for a unit of the durable good that is v periods old at the beginning of period 0. These depreciation rates can be defined recursively, starting with the period 0 depreciation rate for a brand new unit, δ_0^0, using the period 0 vintage asset prices P_v^0 as follows:

$$(15)\ 1 - \delta_0^0 \equiv P_1^0/P^0 .$$

Once δ_0^0 has been defined by (15), the period 0 cross sectional depreciation rate for a unit of the durable good that is one period old at the beginning of period 0, δ_1^0, can be defined using the following equation:

$$(16)\ (1 - \delta_1^0)(1 - \delta_0^0) \equiv P_2^0/P^0 .$$

Note that P_2^0 is the beginning of period 0 asset price of a unit of the durable good that is 2 periods old and it is compared to the price of a brand new unit of the durable, P^0 (which is equal to P_0^0 using the vintage good notation).

Given that the period 0 cross sectional depreciation rates for units of the durable that are 0, 1, 2,…, v − 1 periods old at the beginning of period 0 are defined (these are the depreciation rates δ_0^0, δ_1^0, δ_2^0,…, δ_{v-1}^0), then the period 0 cross sectional depreciation rate for units of the durable that are v periods old at the beginning of period 0 can be defined using the following equation:

$$(17)\ (1 - \delta_v^0) \ldots (1 - \delta_1^0)(1 - \delta_0^0) \equiv P_{v+1}^0/P^0 .$$

It should be clear how the sequence of period 0 vintage asset prices P_v^0 can be converted into a sequence of period 0 vintage depreciation rates. It should also be clear that the sequence of equations (15)-(17) can be repeated using the vintage asset price data pertaining to the beginning of period t, P_v^t, in order to obtain a sequence of period t vintage depreciation rates, δ_v^t. In the depreciation literature, it is usually assumed that the sequence of vintage depreciation rates, δ_v^t, is independent of the period t so that:

$$(18)\ \delta_v^t = \delta_v \quad \text{for all periods t and all vintages v.}$$

The above material shows how the sequence of vintage or used durable goods prices at a point in time can be used in order to estimate depreciation rates. This type of methodology, with a few extra modifications to account for differing ages of retirement, was pioneered by Beidelman (1973) (1976) and Hulten and Wykoff (1981a) (1981b) (1996).[56]

Recall the user cost formula for a new unit of the durable good under consideration defined by (1) above. The same approach can be used in order to define a sequence of period 0 user costs for all vintages v of the durable. Thus suppose that P_{v+1}^{1a} is the *anticipated end of period 0 price* of a unit of the durable good that is v periods old at the beginning of period 0 and let r^0 be the consumer's opportunity cost of capital. Then the discounted to the beginning of period 0 *user cost* of a unit of the durable good that is v periods old at the beginning of period 0, u_v^0, is defined as follows:

[55] Using this notation for vintages, it can be seen that the vintage v = 0 price at the beginning of period t = 0, P_0^0, is equal to the price of a new unit of the good, P^0. If these second hand vintage prices depend on how intensively the durable good has been used in previous periods, then it will be necessary to further classify the durable good not only by its vintage v but also according to the intensity of its use. In this case, think of the sequence of vintage asset prices P_v^0 as corresponding to the prevailing market prices of the various vintages of the good at the beginning of period 0 for a assets that have been used at "average" intensities.

[56] See also Jorgenson (1996) for a review of the empirical literature on the estimation of depreciation rates.

$$(19)\ u_v^0 \equiv P_v^0 - P_{v+1}^{1a}/(1 + r^0) ; \qquad v = 0,1,2, \ldots$$

It is now necessary to specify how the *end* of period 0 anticipated vintage asset prices P_v^{1a} are related to their counterpart *beginning* of period 0 vintage asset prices P_v^0. The assumption that is made now is that the entire sequence of vintage asset prices at the end of period 0 is equal to the corresponding sequence of asset prices at the beginning of period 0 times a general anticipated period 0 inflation rate factor, $(1+i^0)$, where i^0 is the anticipated period 0 (general) asset inflation rate. Thus it is assumed that

$$(20)\ P_v^{1a} = (1 + i^0)P_v^0 ; \qquad v = 1,2, \ldots$$

Substituting (20) and (15)-(18) into (19) leads to the following beginning of *period 0 sequence of vintage user costs*:[57]

$$(21)\ u_v^0 = (1 - \delta_{v-1})(1 - \delta_{v-2}) \ldots (1 - \delta_0)[(1 + r^0) - (1 - \delta_v)(1 + i^0)]P^0/(1 + r^0)$$
$$= (1 - \delta_{v-1})(1 - \delta_{v-2}) \ldots (1 - \delta_0)[r^0 - i^0 + \delta_v(1 + i^0)]P^0/(1 + r^0) ; \qquad v = 0,1,2,\ldots$$

Note that if $v = 0$, then the u_0^0 defined by (21) agrees with the user cost formula for a new purchase of the durable u^0 that was derived earlier in (7).

The sequence of vintage user costs u_v^0 defined by (21) are expressed in terms of prices that are discounted to the *beginning* of period 0. However, as was done in section 4 above, it is also possible to express the user costs in terms of prices that are "discounted" to the *end* of period 0. Thus define the sequence of vintage *end of period 0 user cost* p_v^0 as follows:

$$(22)\ p_v^0 \equiv (1 + r^0)u_v^0 = (1 - \delta_{v-1})(1 - \delta_{v-2}) \ldots (1 - \delta_0)[r^0 - i^0 + \delta_v(1 + i^0)]P^0 ; \quad v = 0,1,2,\ldots$$

If the real interest rate r^{0*} is defined as the nominal interest rate r^0 less the asset inflation rate i^0 and the small terms $\delta_v i^0$ are neglected in (22), then the sequence of end of the period user costs defined by (22) reduces to:

$$(23)\ p_v^0 = (1 - \delta_{v-1})(1 - \delta_{v-2}) \ldots (1 - \delta_0)[r^{0*} + \delta_v]P^0 ; \qquad v = 0,1,2,\ldots$$

Thus if the price statistician has estimates for the vintage depreciation rates δ_v and the real interest rate r^{0*} and is able to collect a sample of prices for new units of the durable good P^0, then the sequence of vintage user costs defined by (23) can be calculated. To complete the model, the price statistician should gather information on the stocks held by the household sector of each vintage of the durable good and then normal index number theory can be applied to these p's and Q's, with the p's being vintage user costs and the Q's being the vintage stocks pertaining to each period. For some worked examples of this methodology under various assumptions about depreciation rates and the calculation of expected asset inflation rates, see Diewert and Lawrence (2000) and Diewert (2003c).[58]

In the following three subsections, the general methodology described above is specialized by making additional assumptions about the form of the vintage depreciation rates δ_v.

6.2 Geometric or Declining Balance Depreciation

[57] When $v = 0$, define $\delta_{-1} \equiv 1$; i.e., the terms in front of the square brackets on the right hand side of (21) are set equal to 1.

[58] Additional examples and discussion can be found in two recent OECD Manuals on productivity measurement and the measurement of capital; see OECD (2001a) (2001b).

The *declining balance method of depreciation* dates back to Matheson (1910; 55) at least.[59] In terms of the algebra presented in section 6.1 above, the method is very simple: all of the cross sectional vintage depreciation rates δ_v^0 defined by (15)-(17) are assumed to be equal to the same rate δ, where δ is a positive number less than one; i.e., for all time periods t and all vintages v, it is assumed that

$$(24)\ \delta_v^t = \delta\ ; \qquad v = 0,1,2,...\ .$$

Substitution of (24) into (22) leads to the following formula for the sequence of *period 0 vintage user costs*:

$$\begin{aligned}(25)\ p_v^0 &= (1-\delta)^v[\, r^0 - i^0 + \delta_v(1+i^0)]P^0\ ; \qquad v = 0,1,2,\ldots \\ &= (1-\delta)^v\, p_0^0\ .\end{aligned}$$

The second set of equations in (25) says *that all of the vintage user costs are proportional to the user cost for a new asset*. This proportionality means that it is not necessary to use an index number formula to aggregate over vintages to form a durable services aggregate. To see this, it is useful to calculate the aggregate value of services yielded by all vintages of the consumer durable at the beginning of period 0. Let q^{-v} be the quantity of the durable purchased by the household sector v periods ago for v = 1,2,…. and let q^0 be the new purchases of the durable during period 0. The beginning of period 0 price for these vintages of age v will be p_v^0 defined by (25) above. Thus the aggregate services of all vintages of the good, including those purchased in period 0, will have the following value, S^0:

$$\begin{aligned}(26)\ S^0 &= p_0^0q^0 + p_1^0q^{-1} + p_2^0q^{-2} + \ldots \\ &= p_0^0q^0 + (1-\delta)\, p_0^0q^{-1} + (1-\delta)^2\, p_0^0q^{-2} + \ldots \qquad \text{using (25)} \\ &= p_0^0\,[q^0 + (1-\delta)q^{-1} + (1-\delta)^2q^{-2} + \ldots\,] \\ &= p_0^0Q^0\end{aligned}$$

where the period 0 aggregate (quality adjusted) quantity of durable services consumed in period 0, Q^0, is defined as

$$(27)\ Q^0 \equiv q^0 + (1-\delta)q^{-1} + (1-\delta)^2q^{-2} + \ldots\ .$$

Thus the *period 0 services quantity aggregate* Q^0 is equal to new purchases of the durable in period 0, q^0, plus one minus the depreciation rate δ times the purchases of the durable in the previous period, q^{-1}, plus the square of one minus the depreciation rate times the purchases of the durable two periods ago, q^{-2}, and so on. The service price that can be applied to this quantity aggregate is p_0^0, the imputed rental price or user cost for a new unit of the durable purchased in period 0.

If the depreciation rate δ and the purchases of the durable in prior periods are known, then the aggregate service quantity Q^0 can readily be calculated using (27). Then using (26), it can be seen that the value of the services of the durable (over all vintages), S^t, decomposes into the price term p_0^0 times the quantity term Q^0. Hence, it is not necessary to use an index number formula to aggregate over vintages using this depreciation model.

6.3 Straight Line Depreciation

[59] A case for attributing the method to Walras (1954; 268-269) could be made but he did not lay out all of the details. Matheson (1910; 91) used the term "diminishing value" to describe the method. Hotelling (1925; 350) used the term "the reducing balance method" while Canning (1929; 276) used the term the "declining balance formula".

Another very common model of depreciation is the *straight line model*.[60] In this model, a most probable length of life for the durable is somehow determined, say L periods, so that after being used for L periods, the durable is scrapped. In the straight line depreciation model, it is assumed that the period 0 cross sectional vintage asset prices P_v^0 follow the following pattern of linear decline relative to the period 0 price of a new asset P^0:

$$(28)\ P_v^0/P^0 = [L - v]/L \qquad \text{for } v = 0, 1, 2, \ldots, L-1.$$

For v = L, L+1, …., it is assumed that $P_v^0 = 0$. Now substitute (20) and (28) into the beginning of the period user cost formula (19) in order to obtain the following sequence of *period 0 vintage user costs* for the durable:

$$\begin{aligned}(29)\ u_v^0 &= P_v^0 - (1 + i^0)P_{v+1}^0/(1 + r^0) \qquad \text{for } v = 0, 1, 2, \ldots, L-1\\ &= [1/L][(L - v) - (L - v - 1)\{(1+i^0)/(1+r^0)\}]P^0\\ &= [(L - v)r^{0*} + 1]\ P^0/L(1 + r^{0*})\end{aligned}$$

where the *asset specific real interest rate for period 0*, r^{0*}, is defined by

$$(30)\ 1 + r^{0*} \equiv (1 + r^0)/(1 + i^0)\ .$$

The user costs for units of the durable good that are older than L periods are zero; i.e., $u_v^0 \equiv 0$ for v ≥ L. Looking at the terms in square brackets on the right hand side of (29), it can be seen that the first term is a real interest opportunity cost for holding and using the unit of the durable that is v periods old (and this imputed interest cost declines as the durable good ages) and the second term is a depreciation term that is equal to the constant rate 1/L .

In this model of depreciation, it is necessary to keep track of household purchases of the durable for L periods and weight up each vintage quantity q^{-v} of these purchases by the corresponding vintage user cost u_v^0 defined by (29) or the end of period vintage user costs p_v^0 defined as $(1+r^0)u_v^0$ could be used.[61]

6.4 One Hoss Shay or Light Bulb Depreciation

The final model of depreciation that is in common use is the "light bulb" or *one hoss shay model of depreciation*.[62] In this model, the durable delivers the *same* services for each vintage: a chair is a chair, no matter what its age is (until it falls to pieces and is scrapped). Thus this model also requires an estimate of the most probable life L of the consumer durable.[63] In this model, it is

---

[60] This model of depreciation dates back to the late 1800's; see Matheson (1910; 55), Garcke and Fells (1893; 98) or Canning (1929; 265-266).

[61] A worked example using this model of depreciation can be found in Diewert (2003b)

[62] This model can be traced back to Böhm-Bawerk (1891; 342). For a more comprehensive exposition, see Hulten (1990; 124) or Diewert (2003b).

[63] The assumption of a single life L for a durable can be relaxed using a methodology due to Hulten: "We have thus far taken the date of retirement T to be the same for all assets in a given cohort (all assets put in place in a given year). However, there is no reason for this to be true, and the theory is readily extended to allow for different retirement dates. A given cohort can be broken into components, or subcohorts, according to date of retirement and a separate T assigned to each. Each subcohort can then be characterized by its own efficiency sequence, which depends among other things on the subcohort's useful life T_i." Charles R. Hulten (1990; 125).

assumed that the sequence of vintage beginning of the period user costs u_v^0 defined by the first line of (29) is *constant* for all vintages younger than the asset lifetime L; i.e., it is assumed that

$$(31)\ u^0 = u_v^0 = P_v^0 - (1 + i^0)P_{v+1}^0/(1 + r^0) \qquad \text{for } v = 0, 1, 2, \ldots, L-1$$
$$= P_v^0 - \gamma P_{v+1}^0$$

where the *discount factor* γ is defined as

$$(32)\ \gamma \equiv (1 + i^0)/(1 + r^0) = 1/(1 + r^{0}*)$$

and the asset specific real interest rate r^0* was defined earlier by (30). Now the second equation in (31) can be used to express the *vintage v asset price* P_v^0 in terms of the common user cost u^0 and the vintage v+1 asset price, P_{v+1}^0, so that

$$(33)\ P_v^0 = u^0 + \gamma P_{v+1}^0 .$$

Now start out using equation (33) with v = 0, then substitute out P_1^0 using (33) with v = 1, then substitute out P_2^0 using (33) with v = 2, etc. until finally the process ends after L such substitutions when P_L^0 is reached and of course, P_L^0 equals zero. The following equation is obtained:

$$(34)\ P^0 = u^0 + \gamma u^0 + \gamma^2 u^0 + \ldots + \gamma^{L-1} u^0$$
$$= u^0\,[1 + \gamma + \gamma^2 + \ldots + \gamma^{L-1}]$$
$$= \{u^0/(1-\gamma)\} - \{u^0\,\gamma^L/(1-\gamma)\} \qquad \text{provided that } \gamma < 1$$
$$= u^0\,(1-\gamma^L)/(1-\gamma) .$$

Now use the last equation in (34) in order to solve for the constant over vintages (beginning of the period) *user cost* for this model, u^0, in terms of the period 0 price for a new unit of the durable, P^0, and the discount factor γ defined by (32):

$$(35)\ u^0 = (1-\gamma)P^0/(1-\gamma^L) .$$

The *end of period 0 user cost*, p^0, is as usual, equal to the beginning of the period 0 user cost, u^0, times the nominal interest rate factor, $1+r^0$:

$$(36)\ p^0 \equiv (1 + r^0)u^0 .$$

The *aggregate services of all vintages* of the good, including those purchased in period 0, will have the following value, S^0:

$$(37)\ S^0 = p_0^0 q^0 + p_1^0 q^{-1} + p_2^0 q^{-2} + \ldots + p_2^0 q^{-(L-1)}]$$
$$= p^0\,[q^0 + q^{-1} + q^{-2} + \ldots + q^{-(L-1)}]$$
$$= p^0 Q^0$$

where the *period 0 aggregate (quality adjusted) quantity of durable services* consumed in period 0, Q^0, is defined as follows for this one hoss shay depreciation model:

$$(38)\ Q^0 \equiv q^0 + q^{-1} + q^{-2} + \ldots + q^{-(L-1)} .$$

Thus in this model of depreciation, the vintage quantity aggregate is the simple sum of household purchases over the last L periods. As was the case with the geometric depreciation model, the one hoss shay model does not require index number aggregation over vintages: there is a constant service price p^0 and the associated period 0 quantity Q^0 is a weighted sum of past purchases for the geometric model and a simple sum over the purchases of the last L periods for the light bulb model.[64]

6.5 The Empirical Estimation of Depreciation Rates

How can the different models of depreciation be distinguished empirically? For durable goods that do not change in quality over time, there are *three possible methods* for determining the sequence of vintage depreciation rates:[65]

- By making a rough estimate of the average length of life L for the durable good and then by *assuming* a depreciation model that seems most appropriate.[66]
- By using cross sectional information on used durable prices at a single point in time and then using equations (15)-(17) above to determine the corresponding sequence of vintage depreciation rates.
- By using cross sectional information on the rental or leasing prices of the durable as a function of the age of the durable and then equations (21) or (22), along with information on the appropriate nominal interest rate and expected durables inflation rate, can be used to determine the corresponding sequence of vintage depreciation rates.

In practice, the third method listed above has not been used (except for rental housing) because the rental markets do not exist or due to difficulties in obtaining the required information on rents by the age of the asset.

Typically, the second method for determining depreciation rates is also not used as described above due to missing information; i.e., not all vintages of the durable are sold on the marketplace at any one point in time. Under these circumstances, an econometric model is constructed that makes use of the limited information on used durable prices but allows the econometrician to estimate the vintage depreciation rates.[67]

7. Unique Durable Goods and the User Cost Approach

In the previous sections, it was assumed that a newly produced unit of the durable good remained the same from period to period. This means that the various vintages of the durable good repeat themselves going from period to period and hence a particular vintage of the good in the current period can be compared with the same vintage in the next period. In particular, consider the period 0 user cost of a new unit of a durable good p_0^0 defined earlier by (8). For convenience, the formula is repeated here:

(39) $p_0^0 = [(1 + r^0) - (1 - \delta_0)(1 + i^0)]P^0 = [r^0 - i^0 + \delta_0(1 + i^0)]P^0$.

Recall that P^0 is the beginning of period 0 purchase price for the durable, r^0 is the nominal opportunity cost of capital that the household faces in period 0, i^0 is the anticipated period 0 inflation rate for the durable good and δ_0 is the one period depreciation rate for a new unit of the durable good. In previous sections, it was assumed that the period 0 user cost p_0^0 for a new unit of

[64] Thus (38) is the quantity aggregate counterpart to (27).
[65] These three classes of methods were noted in Malpezzi, Ozanne and Thibodeau (1987; 373-375) in the housing context.
[66] A length of life L is usually converted into an equivalent geometric depreciation rate δ by setting δ equal to a number between 1/L and 2/L.
[67] See Hall (1971), Beidelman (1973) (1976) and Hulten and Wykoff (1981a) (1981b). See also the discussion of alternative methods for estimating housing depreciation in Malpezzi, Ozanne and Thibodeau (1987; 373).

the durable could be compared with the corresponding period 1 user cost p_0^1 for a new unit of the durable purchased in period 1. This period 1 user cost can be defined as follows:

$$(40)\ p_0^1 = [(1 + r^1) - (1 - \delta_0)(1 + i^1)]P^0 = [r^1 - i^1 + \delta_0(1 + i^1)]P^0 .$$

However, many durable goods are produced as *one of a kind* models. For example, a new house may have many features that are specific to that particular house. An exact duplicate of it is unlikely to be built in the following period. Thus if the user cost for the house is constructed for period 0 using formula (39) where the new house price P^0 plays a key role, then since there will not necessarily be a comparable new house price for the same type of unit in period 1, it will not be possible to construct the period 1 user cost for a house of the same type, p_0^1 defined by (40), since the comparable new house price P^1 will not be available.

Recall the notation that was introduced in section 6.1 above where P_v^t was the second hand market price at the beginning of period t of a unit of a durable good that is v periods old. Define δ_v to be the depreciation rate for a unit of the durable good that is v periods old at the beginning of the period under consideration. Using this notation, the user cost of the house (which is now one period old) for period 1, p_1^1 can be defined as follows:

$$(41)\ p_1^1 \equiv (1 + r^1)P_1^1 - (1 - \delta_1)(1 + i^1)P_1^1$$

where P_1^1 is the beginning of period 1 price for the house that is now one period old, r^1 is the nominal opportunity cost of capital that the household faces in period 1, i^1 is the anticipated period 1 inflation rate for the durable good and δ_1 is the one period depreciation rate for a house that is one period old. For a unique durable good, there is no beginning of period 1 price for a new unit of the durable, P^1, but it is natural to impute this price as the potentially observable market price for the used durable, P_1^1, divided by one minus the period 0 depreciation rate, δ_0; i.e., define an imputed period 1 price for a new unit of the unique durable as follows:

$$(42)\ P^1 \equiv P_1^1/(1 - \delta_0) .$$

If (42) is solved for P_1^1 and the solution is substituted into the user cost defined by (41), then the following expression is obtained for p_1^1, *the period 1 user cost of a one period old unique consumer durable*:

$$(43)\ p_1^1 \equiv (1 - \delta_0)[(1 + r^1) - (1 - \delta_1)(1 + i^1)]P^1$$

If it is further assumed that the unique consumer durable follows the geometric model of depreciation, then

$$(44)\ \delta \equiv \delta_0 = \delta_1 .$$

Substituting (44) into (43) and (40) leads to the following relationship between *the imputed rental cost in period 1 of a new unit of the consumer durable*, p_0^1, and *the period 1 user cost of the one period old consumer durable*, p_1^1:

$$(45)\ p_1^0 = p_1^1/(1 - \delta) .$$

Thus in order to obtain an imputed rental price for the unique consumer durable for period 1, p_0^1, that is *comparable* to the period 0 rental price for a new unit of the consumer durable, p_0^0, it is necessary to make a *quality adjustment* to the period 1 rental price for the one period old durable, p_1^1, by dividing this latter price by one minus the one period geometric depreciation rate, δ. This observation has implications for the quality adjustment of observed market rents of houses. Without this type of quality adjustment, observed dwelling unit rents will have a *downward bias*, since the observed rents do not adjust for the gradual lowering of the quality of the unit due to depreciation of the unit.[68]

Note also that in order to obtain an imputed purchase price for the unique consumer durable for period 1, P^1, that is *comparable* to the period 0 purchase price for a new unit of the consumer durable, P^0, it is necessary to make a *quality adjustment* to the period 1 used asset price for the one period old durable, $P_1^{\,1}$, by dividing this latter price by one minus the period 0 depreciation rate, δ_0; recall equation (23.42) above.[69]

This section is concluded with some observations on the difficulties for economic measurement that occur when it is attempted to determine depreciation rates empirically for unique assets. Consider again equation (42), which allows one to express the potentially observable market price of the unique asset at the beginning of period 1, $P_1^{\,1}$, as being equal to $(1-\delta_0)P^1$, where P^1 is a hypothetical period 1 price for a new unit of the unique asset. If it is assumed that this hypothetical period 1 new asset price is equal to the period 0 to 1 inflation rate factor $(1+i^0)$ times the observable period 0 asset price P^0, then the following relationship between the two observable asset prices is obtained:

(46) $P_1^{\,1} = (1 - \delta_0)(1 + i^0)P^0$.

Thus the potentially observable period 1 used asset price $P_1^{\,1}$ is equal to the period 0 new asset price P^0 times the product of two factors: $(1-\delta_0)$, *a quality adjustment factor* that takes into account the effects of aging on the unique asset, and $(1+i^0)$, a period to period *pure price change factor* holding quality constant. The problem with unique assets is that cross sectional information on used asset prices at any point in time is no longer available to enable one to sort out the separate effects of these two factors. Thus there is a fundamental identification problem with unique assets; without extra information or assumptions, *it will be impossible to distinguish the separate effects of asset deterioration and asset inflation*.[70] In practice, this identification problem is solved by making somewhat arbitrary assumptions about the form of depreciation that the asset is expected to experience.[71]

---

[68] There is an exception to this general observation: if housing depreciation is of the one hoss shay type, then there is no need to quality adjust observed rents for the same unit over time. However, one hoss shay depreciation is empirically unlikely in the housing market since renters are generally willing to pay a rent premium for a new unit over an older unit of the same type. For empirical evidence of this age premium, see Malpezzi, Ozanne and Thibodeau (1987; 378) and Hoffman and Kurz (2002; 19).

[69] This type of quality adjustment to the asset prices for unique consumer durables will always be necessary; i.e., there is no exception to this rule as was the case for one hoss shay depreciation in the context of quality adjusting rental prices.

[70] Special cases of this fundamental identification problem have been noted in the context of various econometric housing models: "For some purposes one might want to adjust the price index for depreciation. Unfortunately, a depreciation adjustment cannot be readily estimated along with the price index using our regression method. … In applying our method, therefore, additional information would be needed in order to adjust the price index for depreciation." Martin J. Bailey, Richard F. Muth and Hugh O. Nourse (1963; 936). "The price index and depreciation are perfectly collinear, so if one cares about the price index, it is necessary to use external information on the geometric depreciation rate of houses." Raymond B. Palmquist (2003; 43).

[71] For example, if the unique asset is a painting by a master, then the depreciation rate can be assumed to be very close to zero. As another example, a reasonable guess at the likely length of life L of the unique asset could be made and then the one hoss shay or straight line depreciation models could be implemented. Alternatively, the length of life L could be converted into an equivalent geometric depreciation rate δ using the conversion rule $\delta = n/L$ where n is a number between 1 and 2.

Housing is the primary example of a unique asset. But in addition to the problems outlined in this section, there are other major problems associated with this particular form of unique asset. These problems will be discussed in the following sections.

8. The User Cost of Owner Occupied Housing

Owner occupied housing is typically an example of a *unique* consumer durable so that the material on the quality adjustment of both stock and rental prices developed in the previous section applies to this commodity. However, owner occupied housing is also an example of a *composite* good; i.e., two distinct commodities are bundled together and sold (or rented) at a single price. The two distinct commodities are:

- the structure and
- the land that the structure sits on.

To model this situation, consider a particular newly constructed dwelling unit that is purchased at the beginning of period 0. Suppose that the purchase price is V^0. This value can be regarded as the sum of a cost of producing the structure, $P_S^0Q_S^0$, where Q_S^0 is the number of square meters of floor space in the structure and P_S^0 is the beginning of period 0 price of construction per square meter, and the cost of the land, $P_L^0Q_L^0$, where Q_L^0 is the number of square meters of the land that the structure sits on and the associated yard and P_L^0 is the beginning of period 0 price of the land per square meter.[72] Thus at the beginning of period 0, *the value of the dwelling unit* is V^0 defined as follows:

$$(47)\ V^0 = P_S^0Q_S^0 + P_L^0Q_L^0 .$$

Suppose that the anticipated price of a unit of a new structure at the beginning of period 1 is P_S^{1a} and that the anticipated price of a unit of land at the beginning of period 1 is P_L^{1a}. Define the *period 0 anticipated inflation rates for new structures and land*, i_S^0 and i_L^0 respectively, as follows:

$$(48)\ 1 + i_S^0 \equiv P_S^{1a}/P_S^0 ;$$
$$(49)\ 1 + i_L^0 \equiv P_L^{1a}/P_L^0 .$$

Let δ_0 be the period 0 depreciation rate for the structure. Then the anticipated beginning of period 1 value for the structure and the associated land is equal to

$$(50)\ V^{1a} = P_S^{1a}(1 - \delta_0)Q_S^0 + P_L^{1a}Q_L^0 .$$

Note the presence of the depreciation term $(1-\delta_0)$ on the right hand side of (50). Should this term be associated with the expected beginning of period 1 price for a new unit of structures P_S^{1a} or with the structures quantity term Q_S^0 ? On the principle that like should be compared to like for prices, it seems preferable to associate $(1-\delta_0)$ with the quantity term Q_S^0. This is consistent with the treatment of unique assets that was suggested in the previous section; i.e., the initial quantity of structures Q_S^0 should be quality adjusted downwards to the amount $(1-\delta_0)\ Q_S^0$ at the beginning of period 1.

Now calculate the cost (including the imputed opportunity cost of capital r^0) of buying the dwelling unit at the beginning of period 0 and (hypothetically) selling it at the end of period 0. The following *end of period 0 user cost or imputed rental cost* R^0 for the dwelling unit is obtained using (47)-(50):

---

[72] If the dwelling unit is part of a multiple unit structure, then the land associated with it will be the appropriate share of the total land space.

(51) $R^0 \equiv V^0(1 + r^0) - V^{1a}$
$= [P_S{}^0Q_S{}^0 + P_L{}^0Q_L{}^0](1 + r^0) - [P_S{}^{1a}(1 - \delta_0)Q_S{}^0 + P_L{}^{1a}Q_L{}^0]$
$= [P_S{}^0Q_S{}^0 + P_L{}^0Q_L{}^0](1 + r^0) - [P_S{}^0 (1 + i_S{}^0)(1 - \delta_0)Q_S{}^0 + P_L{}^0 (1 + i_L{}^0)Q_L{}^0]$
$= p_S{}^0Q_S{}^0 + p_L{}^0Q_L{}^0$

where separate period 0 *user costs of structures and land*, $p_S{}^0$ and $p_L{}^0$, are defined as follows:

(52) $p_S{}^0 = [(1 + r^0) - (1 + i_S{}^0)(1 - \delta_0)]\ P_S{}^0 = [r^0 - i_S{}^0 + \delta_0(1 + i_S{}^0)]\ P_S{}^0$;
(53) $p_L{}^0 = [(1 + r^0) - (1 + i_L{}^0)]\ P_L{}^0 \qquad = [r^0 - i_L{}^0]P_L{}^0$.

Note that the above algebra indicates some of the major determinants of market rents for rental properties. The user cost formulae defined by (52) and (53) can be further simplified if the same approximations that were made in section 4 above are made here (recall equation (9) above); i.e., assume that the terms $r^0 - i_S{}^0$ and $r^0 - i_L{}^0$ can be approximated by a real interest rate r^{0*} and neglect the small term δ_0 times $i_S{}^0$ in (52). Then the user costs defined by (52) and (53) simplify to:

(54) $p_S{}^0 = [(r^{0*} + \delta_0)]\ P_S{}^0$;
(55) $p_L{}^0 = r^{0*}\ P_L{}^0$.

Thus *the imputed rent for an owner occupied dwelling unit* is made up of three main costs:

- The real opportunity cost of the financial capital tied up in the structure;
- The real opportunity cost of the financial capital tied up in the land;
- The depreciation cost of the structure.

The above simplified approach to the user cost of housing can be even further simplified by assuming that the ratio of the quantity of land to structures is fixed and so the aggregate user cost of housing is equal to $[r^{0*} + \delta]P_H{}^0$, where P_H is a quality adjusted housing price index that is based on all properties sold in the country to households during the period under consideration and δ is a geometric depreciation rate that applies to the composite of household structures and land. This super simplified approach is used by Iceland; see Gudnason (2003; 28-29).[73] A variant of this approach is used by the Bureau of Economic Analysis: Lebow and Rudd (2003; 168) note that the US national accounts imputation for the services of owner occupied housing is obtained by applying rent to value ratios for tenant occupied housing to the stock of owner occupied housing. The rent to value ratio can be regarded as an estimate of the applicable real interest rate plus the depreciation rate.[74]

Returning to the period 0 imputed rental cost model for a new structure defined by (47)-(53), now calculate the cost (including the imputed opportunity cost of capital r^1) of buying the used dwelling unit at the beginning of period 1 and (hypothetically) selling it at the end of period 1. Thus at the beginning of period 1, the value of the depreciated dwelling unit is V^1 defined as follows:

[73] The real interest rate that is used is approximately 4% per year and the combined depreciation rate for land and structures is assumed to equal 1.25% per year. The depreciation rate for structures alone is estimated to be 1.5% per year. Property taxes are accounted for separately in the Icelandic CPI. Housing price information is provided by the State Evaluation Board based on property sales data of both new and old housing. The SEB also estimates the value of the housing stock and land in Iceland, using a hedonic regression model based on property sales data. The value of each household's dwelling is collected in the Household Budget Survey.
[74] However, as will be seen in sections 10 and 11 below, this method of imputing the value of Owner Occupied Housing services is likely to give a weight to OOH that is too large.

(56) $V^1 = P_S^1(1 - \delta_0)Q_S^0 + P_L^1Q_L^0$

where P_S^1 is the beginning of period 1 construction price for building a new dwelling unit of the same type and P_L^1 is the beginning of period 1 price of land for the dwelling unit. Note that (56) is an *end of period 0 ex post or actual value* of the dwelling unit whereas the similar expression (50) defined a *beginning of period 0 ex ante or anticipated value* of the dwelling unit.

Suppose that the anticipated price of a unit of a new structure at the beginning of period 2 is P_S^{2a} and that the anticipated price of a unit of land at the beginning of period 2 is P_L^{2a}. Define *the period 1 anticipated inflation rates for new structures and land*, i_S^1 and i_L^1 respectively, as follows:

(57) $1 + i_S^1 \equiv P_S^{2a}/P_S^1$;
(58) $1 + i_L^1 \equiv P_L^{2a}/P_L^1$.

Let δ_1 be the period 1 depreciation rate for the structure. Then *the anticipated beginning of period 2 value for the structure and the associated land* is equal to

(59) $V^{2a} = P_S^{2a}(1 - \delta_0)(1 - \delta_1)Q_S^0 + P_L^{2a}Q_L^0$.

The following end of period 1 *user cost or imputed rental cost* R_1^1 *for a one period old dwelling unit* is obtained using (56)-(59):

(60) $R_1^1 \equiv V^1(1 + r^1) - V^{2a}$
$= [P_S^1(1 - \delta_0)Q_S^0 + P_L^1Q_L^0](1 + r^1) - [P_S^{2a}(1 - \delta_0)(1 - \delta_1)Q_S^0 + P_L^{2a}Q_L^0]$
$= [P_S^1(1 - \delta_0)Q_S^0 + P_L^1Q_L^0](1 + r^1) - [P_S^1(1+i_S^1)(1 - \delta_0)(1 - \delta_1)Q_S^0 + P_L^1(1+i_L^1)Q_L^0]$
$= p_{S1}^1(1- \delta_0)Q_S^0 + p_L^1Q_L^0$

where the period 1 *user costs of one period old structures and land*, p_{S1}^1 and p_L^1, are defined as follows:

(61) $p_{S1}^1 = [(1 + r^1) - (1 + i_S^1)(1 - \delta_1)]\, P_S^1 = [r^1 - i_S^1 + \delta_1(1 + i_S^1)]\, P_S^1$;
(62) $p_L^1 = [(1 + r^1) - (1 + i_L^1)]\, P_L^1 = [r^1 - i_L^1]P_L^1$.

Comparing the period 0 user cost of land p_L^0 defined by (53) with the period 1 user cost of land p_L^1 defined by (62), it can be seen that these user costs have exactly the same form and hence are comparable. However, comparing the period 0 user cost for a new structure p_S^0 defined by (52) with the period 1 user cost for a one period old structure p_{S1}^1 defined by (61), it can be seen that these user costs are not quite comparable unless the period 0 depreciation rate δ_0 is equal to the period 1 depreciation rate δ_1. If declining balance depreciation for structures is assumed, then $\delta_0 = \delta_1 = \delta$, where δ is the common depreciation rate across all periods. Under this assumption, p_{S1}^1 is comparable to the period 0 user cost for a new unit of structures p_S^0. However, even under the assumption of geometric depreciation, it can be seen that the period 1 imputed rent for a one period old dwelling unit R_1^1 defined by (60) is *not* comparable to the corresponding period 0 imputed rent for a new dwelling unit R^0 defined by (51). The imputed rent R^1 that would be comparable to R^0 can be defined as follows:

(63) $R^1 \equiv p_S^1Q_S^0 + p_S^1Q_S^0 = R_1^1 + p_S^1\, \delta\, Q_S^0$

where the period 1 user cost of structures p_S^1 is defined by the right hand side of (61) with δ_1 equal to the common depreciation rate δ and the period 1 user cost of land p_L^1 is defined by (62).

Equation (63) has the following implication for the quality adjustment of the price of a rental property: if R^0 is the observed rent of the unit in period 0 and R_1^{1} is the observed rent for the same dwelling unit in period 1, then the observed rent R_1^{1} is *too low* compared to R^0 and so the period 1 observed rent should be quality adjusted upwards by the period 1 rental price for structures p_S^{1} times the amount of physical depreciation δQ_S^{0} in the structure that occurred in the previous period. This is the same point that was made in section 7 but in this section, the complications due to fact that housing services are a *mixture* of structure and land services are taken into account.

It is evident that the main drivers for the user costs of structures and land are a price index for new dwelling construction, P_S^{t}, and a price index for residential land, P_L^{t}. Most statistical agencies have a constant quality price index for new residential structures, since this index is required in the national accounts in order to deflate investment expenditures on residential structures. This index could be used as an approximation to P_S^{t}.[75] The national accounts also require an imputation for the services of owner occupied housing and thus the constant quality price component of this imputation may be suitable for Consumer Price Index purposes.[76] If the national accounts division also computes quarterly real balance sheets for the economy, then a price index for residential land may be available to the prices division. However, even if this is the case, there will be problems in producing this price index for land on a timely basis and at a monthly frequency.[77] Another possible source of information on land prices may be found in land title registry offices and in the records of real estate firms.

In the following section, the problems involved in obtaining a constant quality price index for either rents or the purchase price of a housing unit are examined in a bit more detail.

9. The Empirical Estimation of Housing Price Indexes

There are two broad approaches to constructing constant quality price indexes for the purchase price of a housing unit:

- The repeat sales approach;
- The hedonic regression approach.

Both of these approaches will be discussed below. The hedonic regression approach can also be applied to the problem of constructing constant quality indexes of rent.

We discuss first the *repeat sales approach*, due to Bailey, Muth and Nourse (1963), who saw their procedure as a generalization of the *chained matched model methodology* that was used by the early pioneers in the construction of real estate price indexes like Wyngarden (1927) and Wenzlick (1952). We first describe this matched model methodology for the case of three periods, which will suffice to illustrate the general case.

---

[75] This index may only be an approximation since it covers the construction of rental properties as well as owner occupied dwellings.

[76] However, the national accounts imputation for the services of Owner Occupied Housing will only be produced on a quarterly basis and so some additional work will be required to produce a price deflator on a monthly basis. Also even though the SNA93 recommends that the imputation for the services of OOH be based on the rental equivalent method, it may be the case that the imputation covers only the imputed depreciation on the structures part of OOH. As was pointed out above, there are two other important additional components that should also be included in OOH services; namely, the imputed real interest on the structures and the land on which the structures sit. These latter two components of imputed expenditures are likely to be considerably larger than the depreciation component.

[77] Another source of information on the value of residential land may be available from the local property tax authorities, particularly if properties are assessed at market values.

Suppose that there is a certain set of housing units S(0,1) that are in scope for the index and are sold in both periods 0 and 1. Denote the sales price for property n sold in period t by V_n^t for $n \in S(0,1)$ and t = 0,1. Let $P^{0,1}$ be the property price index going from period 0 to 1. Then a reasonable stochastic model that relates the ratio of the sales prices of the properties, V_n^1/V_n^0, to the price index $P^{0,1}$ is:

$$(64)\ V_n^1/V_n^0 = P^{0,1} \exp u_n^{0,1}\ ; \qquad n \in S(0,1)$$

where $u_n^{0,1}$ is an independently distributed error term with mean 0 and constant variance. Taking logarithms of both sides of (64) leads to the following linear regression model:

$$(65)\ \ln [V_n^1/V_n^0] = \pi^{0,1} + u_n^{0,1}\ ; \qquad n \in S(0,1)$$

where the single parameter $\pi^{0,1}$ is defined as the logarithm of the price index $P^{0,1}$; i.e.,

$$(66)\ \pi^{0,1} \equiv \ln P^{0,1}.$$

The least squares estimator for $\pi^{0,1}$ is the arithmetic average of the logarithms of the sales price ratios. Exponentiating this estimator leads to the following estimator for the property price index going from period 0 to 1:[78]

$$(67)\ P^{0,1*} \equiv \prod_{n \in S(0,1)} [V_n^1/V_n^0]^{1/N(0,1)}$$

where N(0,1) is the number of houses in the sample that sold in both periods 0 and 1; i.e., it is the number of houses in the set S(0,1). Thus the estimated property price index is simply the equally weighted geometric mean of sales price ratios V_n^1/V_n^0 for all the properties that sold in both periods 0 and 1. This is a typical matched model estimator for an elementary price index.

The above model can be repeated for sales of houses in the target population that sold in both periods 1 and 2. The equations that correspond to (65)-(67) above are (68)-(70) below:

$$(68)\ \ln [V_n^2/V_n^1] = \pi^{1,2} + u_n^{1,2}\ ; \qquad n \in S(1,2)$$

where S(1,2) is the set of houses that sold in both periods 1 and 2 and the parameter $\pi^{1,2}$ is defined as the logarithm of the property price index going from period 1 to 2, $P^{1,2}$; i.e.,

$$(69)\ \pi^{1,2} \equiv \ln P^{1,2}.$$

The least squares estimator for $\pi^{1,2}$ is the arithmetic average of the logarithms of the sales price ratios. Exponentiating this estimator leads to the following estimator for the property price index going from period 1 to 2:

$$(70)\ P^{1,2*} \equiv \prod_{n \in S(1,2)} [V_n^2/V_n^1]^{1/N(1,2)}$$

where N(1,2) is the number of sales of houses in the sample that sold in both periods 1 and 2.

Using the above regression estimates, the *levels* of the property price index, P^t, for periods t = 0,1,2 can be defined as follows:

[78] There will be a (typically small) bias associated with exponentiating an unbiased estimator; i.e., $P^{0,1*}$ will be biased. See Goldberger (1968).

(71) $P^0 \equiv 1$; $P^1 = P^{0,1*}$; $P^2 = P^{0,1*} P^{1,2*}$.

Thus the price index level P^t is set equal to 1 in the base period 0; in period 1, it is equal to the estimated matched model price index going from period 0 to 1, and in period 2, it is equal to the product of the period 1 level times the estimated matched model price index going from period 1 to 2.

The above material explains the chained matched model method that was used prior to the work of Bailey, Muth and Nourse. The innovation made by Bailey, Muth and Nourse (1963) was to reparameterize the regression model defined by (65) and (68) and to add an additional set of estimating equations for repeat sales that took place in periods 0 and 2. Thus the Bailey, Muth and Nourse estimating equations for the case where there are three periods of data on repeat sales are the following ones:

(72) $\ln [V_n^1/V_n^0] = \pi^1 - \pi^0 + u_n^{0,1}$; $n \in S(0,1)$;
(73) $\ln [V_n^2/V_n^1] = \pi^2 - \pi^1 + u_n^{1,2}$; $n \in S(1,2)$;
(74) $\ln [V_n^2/V_n^0] = \pi^2 - \pi^0 + u_n^{0,2}$; $n \in S(0,2)$

where S(0,2) is the set of housing units in the target population that sold in periods 0 and 2 and the π^t are the logarithms of the housing price levels P^t in each period; i.e.,

(75) $\pi^0 \equiv \ln P^0$; $\pi^1 \equiv \ln P^1$; $\pi^2 \equiv \ln P^2$.

It turns out that not all of the parameters π^0, π^1 and π^2 in (72)-(74) can be identified[79] and hence, it is necessary to impose a normalization on the π^t. The natural normalization is

(76) $\pi^0 = 0$ or $P^0 = 1$.

Substituting the normalization (76) into (72)-(74) leads to a simple linear regression model that can be used to obtain least squares estimates for the parameters π^1 and π^2, which we denote by π^{1*} and π^{2*}. Exponentiating these estimates leads to estimates for the period 1 and 2 price levels, P^{1*} and P^{2*} respectively. Hence the Bailey Muth and Nourse estimates for the *housing price levels* in the three periods are defined as follows:

(77) $P^0 \equiv 1$; $P^{1*} \equiv \exp \pi^{1*}$; $P^{2*} \equiv \exp \pi^{2*}$.

It is clear that the Bailey, Muth and Nourse repeat sales model is a big improvement over the original chained repeat sales model since it utilizes the available information on house sales in a statistically more efficient manner.[80]

The above three period model generalizes readily to the general case of T periods considered by Bailey, Muth and Nourse (1963) but the details will be left to the reader to work out.

We now begin our discussion of the hedonic regression approach to constructing constant quality price indexes for housing. *Hedonic regression models* work with *price levels* rather than *price ratios* as dependent variables. Before we discuss a general hedonic regression model for housing, it will be useful to put the above repeat sales model into a levels framework so that the differences between the repeat sales model and a general hedonic model can be assessed.

[79] Adding a constant to each π^t leaves the regression unchanged.
[80] The Bailey, Muth and Nourse regression model has not increased the number of parameters to be estimated (two) but it has added an extra N(0,2) degrees of freedom to the regression model.

Consider the three period case again but now suppose that there is a sample of N houses that sold in each of the three periods. Then a reasonable stochastic model for the prices of the houses V_n^t in each period t might be the following one:

$$(78)\ V_n^t = \alpha_n P^t \exp u_n^t ; \qquad n = 1,2,\ldots,N ; t = 0,1,2$$

where P^t is the housing price index level for period t, α_n is a parameter that reflects the quality of housing unit n relative to "average" quality and u_n^t is an independently distributed error term with mean zero and constant variance. Taking logarithms of both sides of (78) leads to the following system of estimating equations:

$$(79)\ \ln V_n^t = \beta_n + \pi^t + u_n^t ; \qquad n = 1,2,\ldots,N ; t = 0,1,2$$

where the β_n and π^t are defined as follows:

$$(80)\ \beta_n \equiv \ln \alpha_n ; \qquad n = 1,2,\ldots,N;$$
$$(81)\ \pi^t \equiv \ln P^t ; \qquad t = 0,1,2.$$

Equations (79) define a linear regression model in the unknown parameters β_n and π^t. However, as in the previous model, these parameters are not all identified[81] and so a normalization must be imposed. A natural normalization is (76); i.e., we set π^0 equal to 0.

It turns out that the linear regression model defined by (79) and (76) is precisely the same as the country product dummy model (with complete data) for three countries that was invented by Robert Summers (1973) in the context of making price comparisons between countries. It is also a special case of the product dummy hedonic regression model proposed by Aizcorbe, Corrado and Doms (2001).

It is possible to obtain an explicit formula for the least squares estimators for π^1 and π^2 in the linear regression model defined by (76) and (79). The vector of dependent variables in the regression model can be written as the sum of the vectors of exogenous variables times their corresponding least squares estimates plus the vector of least squares residuals. It is well known that the inner product of each exogenous vector with the vector of least squares residuals is zero.[82] This means that the least squares estimators for the unknown parameters in the regression model satisfy the following N+2 equations:

$$(82)\ \sum_{n=1}^N \ln V_n^1 = \sum_{n=1}^N \beta_n^* + N \pi^{1*} ;$$
$$(83)\ \sum_{n=1}^N \ln V_n^2 = \sum_{n=1}^N \beta_n^* + N \pi^{2*} ;$$
$$(84)\ \ln V_n^0 + \ln V_n^1 + \ln V_n^2 = 3\beta_n^* + \pi^{1*} + \pi^{2*} ; \qquad n = 1,2,\ldots,N.$$

Use equations (84) to eliminate the β_n^* from equations (82) and (83) and the resulting two linear equations involving the unknowns π^{1*} and π^{2*} can readily be solved. The solutions are:

$$(85)\ \pi^{1*} = (1/N) \sum_{n=1}^N \ln [V_n^1/V_n^0] ; \pi^{2*} = (1/N) \sum_{n=1}^N \ln [V_n^2/V_n^0].$$

81 A constant can be added to each β_n and subtracted from each π^t without changing the regression model.

82 Write the regression model as $y = X\beta + u$. The vector of least squares estimates β^* for β is defined as $\beta^* \equiv (X^TX)^{-1}X^Ty$. Hence β^* satisfies the system of equations $X^T(X\beta^*) = X^Ty$ or $X^T[y - X\beta^*] = 0_K$ where X has K linearly independent columns. Thus taking the inner product of the kth column of X with $y - X\beta^*$ gives us 0 for each column k.

Using the inverses of equations (81), the π^{t*} defined by equations (85) translate into the following estimates for the period 1 and 2 price levels, P^{1*} and P^{2*} respectively:

(86) $P^{1*} = \prod_{n=1}^{N} [V_n^1/V_n^0]^{1/N}$; $P^{2*} = \prod_{n=1}^{N} [V_n^2/V_n^0]^{1/N}$.

Thus this complete information country product dummy model leads to the geometric mean of the period 1 values relative to the corresponding period 0 values, $\prod_{n=1}^{N} [V_n^1/V_n^0]^{1/N}$, as the estimate for the period 1 housing price level P^{1*}, and to the geometric mean of the period 2 values relative to the corresponding period 0 values, $\prod_{n=1}^{N} [V_n^2/V_n^0]^{1/N}$, as the estimate for the period 2 housing price level P^{2*}. Note that this result is very similar to that of the chained matched model originally proposed by Wyngarden (1927) and Wenzlick (1952), except that instead of using the chain principle, the country product dummy method ends up using the fixed base principle.

We now consider a more realistic model where *not* every house in the sample trades in *each* period. In order to minimize notational complexities, we will consider only the case of two periods. Using our earlier notation, let S(0,1) be the set of housing units that sold in both periods 0 and 1. Taking into account the normalization (76), the estimating equations corresponding to these houses are:

(87) $\ln V_n^0 = \beta_n + u_n^0$; $\quad n \in S(0,1)$;
(88) $\ln V_n^1 = \beta_n + \pi^1 + u_n^1$; $\quad n \in S(0,1)$.

Let S(0~1) denote the set of housing units in the target population that sold in period 0 but not in period 1. The estimating equations for these observations are:

(89) $\ln V_m^0 = \gamma_m + u_m^0$; $\quad m \in S(0\sim1)$

where γ_m is the logarithm of the quality adjustment factor for the mth housing unit that sold in period 0 but not in period 1. Similarly, let S(1~0) denote the set of housing units in the target population that sold in period 1 but not in period 0. The estimating equations for these observations are:

(90) $\ln V_k^1 = \delta_k + u_k^0$; $\quad k \in S(1\sim0)$

where δ_k is the logarithm of the quality adjustment factor for the kth housing unit that sold in period 1 but not in period 0. The linear regression model defined by equations (87)-(90) is the same as the two country version of Summer's (1973) *country product dummy model* (with incomplete information); it is also identical to the two period case of the Aizcorbe, Corrado and Doms (2001) *dummy product hedonic regression model*.

Let π^{1*}, β_n^*, γ_m^* and δ_k^* denote the least squares estimates of the parameters π^1, β_n, γ_m and δ_k that appear in (87)-(90). The stacked vector of dependent variables in equations (87)-(90) can be written as the sum of the vectors of exogenous variables times their corresponding least squares estimates plus the vector of least squares residuals. As noted above, the inner product of each exogenous vector with the vector of least squares residuals is zero. This means that the least squares estimators for the unknown parameters in the regression model satisfy the following equations:[83]

(91) $\sum_{n \in S(0,1)} \ln V_n^1 + \sum_{k \in S(1\sim0)} \ln V_k^1$
$\quad = \sum_{n \in S(0,1)} \beta_n^* + N(0,1)\,\pi^{1*} + \sum_{k \in S(1\sim0)} \delta_k^* + N(1\sim0)\,\pi^{1*}$;
(92) $\ln V_n^0 + \ln V_n^1 = 2\beta_n^* + \pi^{1*}$; $\quad n \in S(0,1)$;

[83] This technique of proof was used by Diewert (2003a) in the context of a hedonic regression model.

(93) $\ln V_m^0 = \gamma_m^*$; $m \in S(0{\sim}1)$;
(94) $\ln V_k^1 = \delta_k^*$; $k \in S(1{\sim}0)$

where N(0,1) is the number of housing units that traded in both periods and N(1~0) is the number of housing units that sold in period 1but not period 0.

Use equations (94) to eliminate the δ_k^* in equation (91) and use equations (92) to eliminate the β_n^* from equation (91). The resulting equation shows that π^{1*} is equal to:

(95) $\pi^{1*} = [1/N(0,1)] \sum_{n \in S(0,1)} \ln [V_n^1/V_n^0]$

which is the arithmetic average of the logarithms of the sales price ratios for the matched models in the two periods. Exponentiating (95) shows that this simple hedonic regression model, where each housing unit has only a single dummy variable characteristic, leads to a period 0 to 1 price index that is equal to the equally weighted geometric mean of the selling prices in period 1 divided by the geometric mean of the corresponding selling prices of the matched models in period 0. Thus *in the two period case, the dummy variable hedonic regression model and the country product dummy model give exactly the same result as the matched model method that is used by statistical agencies*.

We now consider more general hedonic regression models[84]. A more *general hedonic regression model* is the following one:

(96) $\ln V_n^t = \pi^t + \sum_{k=1}^{K} z_{nk}^t \beta_k + u_n^t$; $t = 0,1,2,\ldots,T$; $n \in S(t)$

where S(t) is the set of housing units in the target population that sold (or were rented) in period t. As usual, u_n^t is an independently distributed error term with mean 0 and constant variance, V_n^t is the observed selling price (or rent) of housing unit n in period t and z_{nk}^t is the amount of characteristic k that this housing unit possesses. The parameter π^t is equal to the logarithm of the constant quality price index for period t, P^t, so that $\pi^t = \ln P^t$ for $t = 0,1,\ldots,T$. The unknown parameter β_k transforms amounts of the kth characteristic z_k into constant quality utility units for $k = 1,\ldots,K$.[85]

What are the advantages and disadvantages of the general hedonic regression model defined by (96) compared to the Bailey, Muth and Nourse repeat sales model or to the closely related product dummy hedonic regression models of Summers and Aizcorbe, Corrado and Doms?

The *main advantage* of the general hedonic regression model is that it uses *all* of the information on housing sales in each sample period in a nontrivial way whereas the repeat sales model does not use any information at all on isolated sales that take place in only one of the sample periods. We showed that in the two period case, the product dummy regression model led to an estimator of quality adjusted price change that did not use any information on unmatched sales. This feature of the product dummy hedonic regression model carries over to the case of many time periods; i.e., it is intuitively obvious that when an observation has its very own dummy variable in a linear regression model, then this observation will not be used to determine any of the other parameters in the model. Thus if the unmatched prices in the sample of housing prices behave differently than the matched prices, it can be seen that a general hedonic regression model can generate quite different

[84] The main features of a general hedonic regression model were laid out in Court (1939). This publication was not readily available to researchers and so the technique was not used widely until the work of Griliches (1971a) (1971b) popularized the technique. For a recent survey of the hedonic regression technique for making quality adjustments, see Triplett (2002).

[85] If the vector of characteristics contains a constant term, then there will be exact multicollinearity between this constant term and the time dummy variables, the π^t. Under these conditions, it will be necessary to make a normalization on the parameters such as $\pi^0 = 0$.

price indexes than models that rely only on matched prices.[86] Put another way, a general hedonic regression model uses *all* of the sample data in a nontrivial way and not just the data that can be matched.

The *main disadvantage* of the general hedonic regression model compared to the repeat sales model and the product dummy hedonic regression model is the difficulty in determining just which characteristics should be included in the model. A closely related issue is that it is difficult to determine what the appropriate functional form for the hedonic regression is.[87] In section 10.4 below, some of the functional form problems associated with hedonic regression models for housing will be discussed in more detail.[88] But in general, hedonic regression models suffer from a lack of *reproducibility*; i.e., different statisticians and econometricians will collect data on different characteristics of housing and assume different functional forms for the hedonic regression model (96) and thus come up with different measures of quality adjusted price change.

We summarize the advantages and disadvantages of the repeat sales model as compared to a general hedonic regression model as follows.

The *main advantage* of the repeat sales model is:

- Reproducibility; i.e., different statisticians given the same data on the sales of housing units will come up with the same estimate of quality adjusted price change.

The *main disadvantages* of the repeat sales model are:

- It does not use all of the available information on housing unit sales; it uses only information on housing units that have sold more than once during the sample period.
- It cannot deal adequately with depreciation of the housing structure; i.e., depreciation is exactly collinear with the time dummy variables π^t and thus cannot be distinguished from the effects of price change. Conversely, a general hedonic regression model for housing can adjust for the effects of depreciation if the age of the structure is known at the time of sale (or rental).
- It cannot deal adequately with housing units that have undergone major repairs or renovations.[89] Conversely, a general hedonic regression model for housing can adjust for the effects of renovations and extensions if (real) expenditures on renovations and extensions are known at the time of sale (or rental).[90]

---

[86] For empirical evidence on this point, see the hedonic regression studies of Silver and Heravi (2001) (2002) (2003).
[87] Functional form problems for hedonic regressions are discussed in Diewert (2003a) (2003c).
[88] In particular, many hedonic regression studies use the logarithm of a transaction price as the dependent variable. This specification of the hedonic model is usually not consistent with the additive nature of the structure and land components of a property and the multiplicative nature of the depreciation adjustment as appears in equations (47) and (56) which defined the value of a specific property in successive periods.

[89] Case and Shiller (1989) use a variant of the repeat sales method using US data on house sales in four major cities over the years 1970-1986. They attempt to deal with the depreciation and renovation problems as follows: "The tapes contain actual sales prices and other information about the homes. We extracted from the tapes for each city a file of data on houses sold twice for which there was no apparent quality change and for which conventional mortgages applied." Karl E. Case and Robert J. Shiller (1989; 125-126).
[90] However, usually information on maintenance and renovation expenditures is not available in the context of estimating a hedonic regression model for housing. Malpezzi, Ozanne and Thibodeau (1987;375-6) comment on this problem as follows: "If all units are identically constructed, inflation is absent, and the rate of maintenance and repair expenditures is the same for all units, then precise measurement of the rate of depreciation is possible by observing the value or rent of two or more units of different ages. … To accurately estimate the effects of aging on values and rents, it is necessary to control for inflation, quality differences in housing units, and location. The hedonic technique controls

Both the repeat sales and general hedonic regression approaches to the construction of constant quality price indexes for housing suffer from another problem that has not been mentioned up to now and that is that both of these methods do not allow the prices to be weighted according to their economic importance. Thus if the statistical agency adopts a superlative index[91] as its target index, then prices should be weighted by either quantities or by expenditure shares and from this perspective, equally weighted geometric means of price relatives are not necessarily close to their superlative counterparts. The regression models discussed in this section have not made any mention of weights so the resulting measures of quality adjusted price change could be different from their weighted counterparts.[92] Another problem that has not been discussed is the possibility that house sales prices might exhibit seasonal fluctuations.[93] The general hedonic regression model cold accommodate seasonal prices by having seasonal dummy variables as explanatory variables.

Our conclusion at this point is that there is no completely satisfactory solution to the problems involved in constructing constant quality price indexes for the stock of owner occupied housing. The hedonic regression approach seems to be superior in principle to the repeat sales approach since the latter approach cannot deal adequately with depreciation and renovations to the structure part of a housing unit. However, in practice, the hedonic regression approach has limitations due to its lack of reproducibility and the lack of information on repairs and renovations.

There are many other difficulties associated with measuring the price and quantity of Owner Occupied Housing services. The following section discusses some of the problems involved in modeling the costs of certain expenditures that are tied to the ownership of a home.

10. The Treatment of Costs Tied to Owner Occupied Housing

There are many costs that are quite directly tied to home ownership. However, it is not always clear how these costs can be decomposed into price and quantity components. Several of these cost components are listed below and some suggestions for forming their associated prices are suggested.

10.1 The Treatment of Mortgage Interest Costs

The derivation of the user cost or expected rental price that an owner of a home should charge for the use of the dwelling unit for one period implicitly assumed that the owner had no mortgage interest costs so that the interest rate r^0 referred to the owner's opportunity cost of equity capital. In this section, the case where the owner has a mortgage on the property is considered.

Recall the notation in the previous section where the user cost or imputed rental cost, R^0, for an equity financed dwelling unit was obtained; see (51). Suppose now that the property purchase is partly financed by a mortgage of M^0 dollars at the beginning of period 0. Let f^0 be the fraction of the beginning of period 0 market value of the property that is financed by the mortgage so that

(97) $M^0 = f^0V^0 = f^0[P_S{}^0Q_S{}^0 + P_L{}^0Q_L{}^0]$.

for differences in dwelling quality and inflation rates but cannot control for most differences in maintenance (except to the extent that they are correlated with location)."

[91] Superlative indexes were initially introduced as approximations to economic cost of living indexes; see Diewert (1976) (1978). But it turns out that various superlative indexes emerge as useful target indexes from the perspectives of the fixed basket, axiomatic and stochastic approaches as well; see Diewert (2002a; 565-581).

[92] Diewert (2002b) (2003c) discusses how weights can be introduced into both the product dummy and general hedonic regression models.

[93] Case and Shiller (1989; 127) note that US house prices tend to have a seasonal peak in July.

Let the one period nominal mortgage interest rate be r_M^0. The owner's period 0 benefits of owning the dwelling unit remain the same as in section 8 and are equal to V^{1a} defined by (50). However, the period 0 costs are now made up of an explicit mortgage interest cost equal to $M^0(1+r_M^0)$ plus an imputed equity cost equal to $(1-f^0)V^0(1+r^0)$. Thus the new imputed rent for using the property during period 0 is now

$$
\begin{aligned}
(98)\ R^0 &\equiv (1 - f^0)V^0(1 + r^0) + M^0(1 + r_M^0) - V^{1a} \\
&= (1 - f^0)[P_S^0Q_S^0 + P_L^0Q_L^0](1 + r^0) + f^0[P_S^0Q_S^0 + P_L^0Q_L^0](1 + r_M^0) \\
&\quad - [P_S^{1a}(1 - \delta_0)Q_S^0 + P_L^{1a}Q_L^0] \\
&= p_S^{0*}Q_S^0 + p_L^{0*}Q_L^0
\end{aligned}
$$

where the new mortgage interest adjusted period 0 *user costs of structures and land*, p_S^{0*} and p_L^{0*}, are defined as follows:

$$
\begin{aligned}
(99)\quad p_S^{0*} &\equiv [(1 + r^0)(1 - f^0) + (1 + r_M^0)f^0 - (1 + i_S^0)(1 - \delta_0)]P_S^0 \\
&= [(r^0 - i_S^0)(1 - f^0) + (r_M^0 - i_S^0)f^0 + \delta_0(1 + i_S^0)]P_S^0 ;
\end{aligned}
$$

$$
\begin{aligned}
(100)\ p_L^{0*} &\equiv [(1 + r^0)(1 - f^0) + (1 + r_M^0)f^0 - (1 + i_L^0)]P_S^0 \\
&= [(r^0 - i_L^0)(1 - f^0) + (r_M^0 - i_L^0)f^0]P_S^0 .
\end{aligned}
$$

Comparing the new user costs for structures and land defined by (99) and (100) with the corresponding equity financed user costs defined by (52) and (53) in the previous section, it can be seen that the old equity opportunity cost of capital r^0 is now replaced by a weighted average of this equity opportunity cost and the mortgage interest rate, $r^0(1-f^0) + r_M^0f^0$, where f^0 is the fraction of the beginning of period 0 value of the dwelling unit that is financed by the mortgage.

Central bankers often object to the inclusion of mortgage interest in a Consumer Price Index. However, examination of the last equation in (99) and in (100) shows that the *nominal* mortgage interest rate r_M^0 has an offsetting benefit due to *anticipated price inflation* in the price of structures, i_S^0 in (99), and in the price of land, i_L^0 in (100), so as usual, what counts in these user cost formulae are *real* interest costs rather than *nominal* ones.

10.2 The Treatment of Property Taxes

Recall the user costs of structures and land defined by (52) and (53) in section 8 above. It is now supposed that the owner of the housing unit must pay the property taxes T_S^0 and T_L^0 for the use of the structure and land respectively during period 0.[94] Define *the period 0 structures tax rate* τ_S^0 and *land tax rate* τ_L^0 as follows:

$$(101)\ \tau_S^0 \equiv T_S^0/P_S^0Q_S^0 ;$$
$$(102)\ \tau_L^0 \equiv T_L^0/P_L^0Q_L^0 .$$

The *new imputed rent for using the property during period 0*, R^0, including the property tax costs, is defined as follows:

$$
\begin{aligned}
(103)\ R^0 &\equiv V^0(1 + r^0) + T_S^0 + T_L^0 - V^{1a} \\
&= [P_S^0Q_S^0 + P_L^0Q_L^0](1 + r^0) + \tau_S^0P_S^0Q_S^0 + \tau_L^0P_L^0Q_L^0 \\
&\quad - [P_S^0(1 + i_S^0)(1 - \delta_0)Q_S^0 + P_L^0(1 + i_L^0)Q_L^0] \\
&= p_S^0Q_S^0 + p_L^0Q_L^0
\end{aligned}
$$

[94] If there is no breakdown of the property taxes into structures and land components, then just impute the overall tax into structures and land components based on the beginning of the period values of both components.

where separate period 0 *tax adjusted user costs of structures and land*, p_S^0 and p_L^0, are defined as follows:

$$(104)\ p_S^0 \equiv [(1 + r^0) - (1 + i_S^0)(1 - \delta_0) + \tau_S^0]P_S^0$$
$$= [r^0 - i_S^0 + \delta_0(1 + i_S^0) + \tau_S^0]P_S^0\ ;$$

$$(105)\ p_L^0 \equiv [(1 + r^0) - (1 + i_L^0) + \tau_L^0]P_L^0$$
$$= [r^0 - i_L^0 + \tau_S^0]P_L^0\ .$$

Thus the property tax rates, τ_S^0 and τ_L^0 defined by (101) and (102), enter the user costs of structures and land, p_S^0 and p_L^0 defined by (104) and (105), in a simple additive manner; i.e., these terms are additive to the previous depreciation and real interest rate terms.[95]

10.3 The Treatment of Property Insurance

At first glance, it would seem that *property insurance* could be treated in the same manner as the treatment of property taxes in the previous subsection. Thus let C_S^0 be the cost of insuring the structure at the beginning of period 0 and define *the period 0 structures premium rate* γ_S^0 as follows:

$$(106)\ \gamma_S^0 \equiv C_S^0/P_S^0Q_S^0\ .$$

The new *imputed rent* for using the property during period 0, R^0, including property tax and insurance costs, is defined as follows:

$$(107)\ R^0 \equiv V^0(1 + r^0) + T_S^0 + T_L^0 + C_S^0 - V^{1a}$$
$$= [P_S^0Q_S^0 + P_L^0Q_L^0](1 + r^0) + \tau_S^0P_S^0Q_S^0 + \tau_L^0P_L^0Q_L^0 + \gamma_S^0\, P_S^0Q_S^0$$
$$- [P_S^0(1 + i_S^0)(1 - \delta_0)Q_S^0 + P_L^0(1 + i_L^0)Q_L^0]$$
$$= p_S^0Q_S^0 + p_L^0Q_L^0$$

where separate period 0 *tax and insurance adjusted user costs of structures and land*, p_S^0 and p_L^0, are defined as follows:

$$(108)\ p_S^0 \equiv [(1 + r^0) - (1 + i_S^0)(1 - \delta_0) + \tau_S^0 + \gamma_S^0]P_S^0$$
$$= [r^0 - i_S^0 + \delta_0(1 + i_S^0) + \tau_S^0 + \gamma_S^0]P_S^0\ ;$$

$$(109)\ p_L^0 \equiv [(1 + r^0) - (1 + i_L^0) + \tau_L^0]P_L^0$$
$$= [r^0 - i_L^0 + \tau_S^0]P_L^0\ .$$

Thus the insurance premium rate γ_S^0 appears in the user cost of structures, p_S^0 defined by (108), in an additive manner, analogous to the additive property tax rate term.[96] If it is desired to have a separate CPI price component for insurance, then the corresponding period 0 and 1 prices can be defined as $\gamma_S^0P_S^0$ and $\gamma_S^1P_S^1$ respectively while the corresponding period 0 and 1 expenditures can

[95] If the price statistician uses the national accounts imputation for the value of owner occupied housing services, care should be taken to ensure that the value of property taxes is included in this imputation.
[96] This treatment of property insurance dates back to Walras (1954; 268-269).

be defined as $\gamma_S{}^0P_S{}^0Q_S{}^0$ and $\gamma_S{}^1P_S{}^1(1-\delta)Q_S{}^0$ respectively.[97] Of course, if this separate treatment is implemented, then these terms have to be dropped from the corresponding user costs of structures.

The above treatment of property taxation and insurance assumes that the property taxes and the premium payments are made at the *end* of the period under consideration; see (107) above. While this may be an acceptable approximation for the payment of property taxes, it is not acceptable for the payment of insurance premiums: the premium *must* be paid at the *beginning* of the period of protection rather than at the end. When this complication is taken into account, the user cost of structures becomes

$$(110)\ p_S{}^0 \equiv [(1 + r^0) - (1 + i_S{}^0)(1 - \delta_0) + \tau_S{}^0 + \gamma_S{}^0(1 + r^0)]P_S{}^0$$
$$= [r^0 - i_S{}^0 + \delta_0(1 + i_S{}^0) + \tau_S{}^0 + \gamma_S{}^0(1 + r^0)]P_S{}^0 .$$

There are some additional problems associated with the modeling of property insurance:

- The above user cost derivations assume that the *risk* of property damage remains constant from period to period. If the risk of damage changes, then an argument can be made for quality adjustment of the premium to hold constant the risk so that like can be compared with like.
- The *gross premium approach* to insurance is taken in the above treatment; i.e., it is assumed that dwelling owners pay premiums for property protection services, no matter whether they have a claim or not. In the *net premium approach*, payments to settle claims are subtracted from the gross premium payments.
- The property protection may not be complete; i.e., the insurance policy may have various limitations on the type of claim that is allowed and there may be a deductible or damage threshold, below which no claim is allowed. If the deductible changes from period to period, then the price statistician is faced with a rather complex quality adjustment problem.

Thus it can be seen that there are many difficult problems that remain to be resolved in this area.

10.4 The Treatment of Maintenance and Renovation Expenditures

Another problem associated with home ownership is the treatment of *maintenance expenditures, major repair expenditures and expenditures associated with renovations or additions*.

Empirical evidence suggests that the normal decline in a structure due to the effects of aging and use can be offset by maintenance and renovation expenditures. How exactly should these expenditures be treated in the context of modeling the costs and benefits of home ownership?

A common approach in the national accounts literature is to treat major renovation and repair expenditures as capital formation and smaller routine maintenance and repair expenditures as current expenditures. If this approach is followed in the CPI context, then these smaller routine maintenance expenditures can be treated in the same manner as other nondurable goods and services. The major renovation and repair expenditures do not enter the CPI in the period that they are made but these expenditures are capitalized and added to expenditures on new structures for the period under consideration, so that period 0 investment in structures in constant dollars, $I_S{}^0$ say[98],

[97] Similarly, if it is desired to have a separate CPI price component for property taxes on structures, then the corresponding period 0 and 1 prices can be defined as $\tau_S{}^0P_S{}^0$ and $\tau_S{}^1P_S{}^1$ respectively while the corresponding period 0 and 1 expenditures can be defined as $\tau_S{}^0P_S{}^0Q_S{}^0$ and $\tau_S{}^1P_S{}^1(1-\delta)Q_S{}^0$ respectively.

[98] Let $VI_S{}^0$ be the nominal value of investment in new owner occupied structures in period 0 plus the value of major renovation expenditures made during period 0. Then the constant dollar quantity of investment could be defined as $I_S{}^0 \equiv VI_S{}^0/P_S{}^0$ where $P_S{}^0$ is the period 0 construction price index for new structures.

would include both types of expenditures. Let Q_S^0 and Q_S^1 be the stocks (in constant quality units) of owner occupied structures in the reference population at the beginning of period 0 and 1 respectively. Then if the geometric model of depreciation is used, so that the constant period to period depreciation rate δ is applicable, then the beginning of period 1 stock of owner occupied structures Q_S^1 is related to the beginning of period 0 stock of structures Q_S^0 and the period 0 investment in structures I_S^0 according to the following equation:

$$(111)\ Q_S^1 = (1-\delta)Q_S^0 + I_S^0 .$$

Thus if declining balance depreciation is assumed for structures, then the treatment of major repair and renovation expenditures does not pose major *conceptual* problems using a conventional capital accumulation model: it is only necessary to have an estimate for the monthly or quarterly depreciation rate δ, a starting value for the stock of owner occupied structures for some period, information on new purchases of residential housing structures by the household sector, information on expenditures by owners on major repairs and renovations and a construction price index for new residential structures. With this information on a timely basis, up to date CPI weights for the stock of owner occupied structures could be constructed.[99]

We now look at how major repair and renovation expenditures could be treated in a repeat sales regression model that used transactions data on the sale of the *same* housing unit in two or more periods. In order to minimize notational complexities, consider a highly simplified situation where data on the sale of N houses of a relatively homogeneous type for two consecutive periods are available. Suppose these *sale prices* are V_n^0 for period 0 and V_n^1 for period 1, for $n = 1,2,\ldots,N$. Suppose that a *price index for structures* of this type of property in period 0, P_S^0, and a corresponding *price index for land* in period 0, P_L^0, have been constructed.[100] The price statistician's problem is to use the data on the matched sales for the two periods in order to construct estimates of these two indices for period 1; i.e., the problem is to construct P_S^1 and P_L^1.

The period 0 *dwelling unit values* for the N properties can be decomposed into the structure and land components as follows:

$$(112)\ V_n^0 = V_{Sn}^0 + V_{Ln}^0 = \alpha_n P_S^0 Q_{Sn}^0 + \beta_n P_L^0 Q_{Ln}^0 ; \qquad n = 1,2,\ldots,N$$

where V_{Sn}^0 and V_{Ln}^0 are the estimated period 0 values of the structure and land of property n in period 0, P_S^0 and P_L^0 are the (known) price index values for structures and land for all properties of this type in period 0 and Q_{Sn}^0 and Q_{Ln}^0 are (known) estimates of the quantity of structures and land for property n. The numbers α_n and β_n are *property n quality adjustment factors* that convert the property standardized values of structures and land, $P_S^0Q_{Sn}^0$ and $P_L^0Q_{Sn}^0$ respectively, into the period 0 actual market values, V_{Sn}^0 and V_{Ln}^0 respectively; i.e., if estimates of the period 0 market values of the structures and land for property n are available, then α_n and β_n can be defined as follows:

$$(113)\ \alpha_n \equiv V_{Sn}^0/P_S^0Q_{Sn}^0 ;\ \ \beta_n \equiv V_{Ln}^0/P_L^0Q_{Ln}^0 ; \qquad n = 1,\ldots,N.$$

Suppose that information on the dollar amount of major repairs and renovations made to property n during period 0, VR_n^0, is also available for each property n in the sample of properties. Then the period 1 value for property n, V_n^1, should be approximately equal to

[99] However, the *practical* problems involved in obtaining all of this information on a timely basis are not trivial. Variants of this approach were used by Christensen and Jorgenson (1969) and Leigh (1980) in order to construct estimates of the stock of residential structures in the US.
[100] If these period 0 indices are not available, then set P_S^0 and P_L^0 equal to 1.

(114) $V_n^1 = \alpha_n P_S^1(1-\delta)Q_{Sn}^0 + VR_n^0 + \beta_n P_L^1 Q_{Ln}^0$; $n = 1,2,\ldots,N$

where δ is the geometric depreciation rate for structures. All of the variables on the right hand side of (114) are assumed to be known with the exception of the period 1 price index values for structures and land, P_S^1 and P_L^1 respectively, and the one period geometric depreciation rate, δ. If the number of observations N is greater than three, then it would appear that these three parameters, P_S^1, P_L^1 and δ, could be estimated by a linear regression using the N equations in (114) as estimating equations. However, it turns out that this is not quite correct. The problem is that the parameters P_S^1 and $(1-\delta)$ appear in (114) in a multiplicative fashion so that while the product of these two terms will be nicely identified by the regression, *the individual terms cannot be uniquely identified*. This is just a reappearance of the same problem that was discussed earlier in section 7 on unique consumer durables: the separate effects of aging of the asset (depreciation or capital consumption) and price appreciation over time *cannot* be separately identified using just market data on resales if the housing unit is regarded as a unique asset.[101]

There are three possible solutions to this identification problem:

- Use an external estimate of the depreciation rate δ;
- Use an external construction price index P_S^1 instead of estimating it as a parameter in equations (114);
- Abandon the repeat sales approach and use a hedonic regression approach instead.

What would a hedonic regression model look like, taking into account the approximate additivity of the value of the housing structure and the value of the land that the structure sits on? If the renovations problem is ignored and geometric depreciation of the structure is assumed, then the value of a housing unit n in period t that is v periods old, V_n^t, should be approximately equal to the depreciated value of the structure plus the value of the land plus an error term; i.e., the following relationship should hold approximately:

(115) $V_n^t = P_S^t(1-\delta)^v Q_{Sn} + P_L^t Q_L + u_n^t$

where δ is the one period geometric depreciation rate, Q_{Sn} is the number of square meters of floor space of the original structure for housing unit n, Q_L is the number of square meters of land that the housing structure sits on and u_n^t is an error term. P_S^t is the beginning of *period t price level for structures* of this type and P_L^t is the corresponding *price of land* for this class of housing units. As long as there is more than one vintage of structure in the sample (i.e., more than one v), then the parameters P_S^t, P_L^t and δ can be identified by running a nonlinear regression model using equations (115). Why can the price levels be identified in the present hedonic regression model whereas they could not be identified in the repeat sales model? The answer is that the hedonic model (115) does *not* assume property specific quality adjustment factors for each housing unit; instead, *all* of the housing units in the sample are assumed to be of comparable quality once prices are adjusted for the age of the unit and the quantity (in square meters) of original structure and the quantity of land.

[101] Recall equation (46) above. This fundamental identification problem was recognized by Bailey, Muth and Nourse (1963; 936) in the original repeat sales housing article but it was ignored by them and subsequent users of the repeat sales methodology. Another problem with the housing hedonic regression literature is that usually, the logarithm of the purchase price is taken as the dependent variable in the regression. While this specification has some advantages, it does not recognize properly the additive nature of the structure and land components of the housing property. A final problem with the traditional hedonic housing literature is that usually, separate price indices for land and structures are not estimated. It is important to allow for separate price indices for these two components since usually, the price of land is more volatile and tends to increase faster than the price of structures over long periods of time.

Unfortunately, many housing structures that may have started their lives as identical structures do not remain the same over time, due to differing standards of maintenance as well as major renovations and additions to some of the structures. To model this phenomenon, let R_n^t be real maintenance, repair and renovation expenditures on housing unit n during period t and suppose that these real expenditures depreciate at the geometric rate δ_R. It is reasonable to assume that these expenditures *add* to the value of the housing unit and so equations (115) should be replaced by the following equations:

$$(116)\ V_n^t = P_S^t(1-\delta)^v Q_{Sn} + P_R^t[R_n^t + (1-\delta_R)R_n^{t-1} + (1-\delta_R)^2 R_n^{t-2} + \ldots + (1-\delta_R)^v R_n^{t-v}] + P_L^t Q_L + u_n^t$$

where P_R^t is the period t price level for real maintenance, repair and renovation expenditures on this class of housing units. If information on these real renovation and repair expenditures, R_n^t, R_n^{t-1}, R_n^{t-2},…,R_n^{t-v}, is available for each housing unit in the sample of housing units that sold in period t, then the parameters P_S^t, P_L^t, P_R^t, δ and δ_R can be identified by running a nonlinear regression model using equations (116).[102]

However, a major practical problem with implementing a hedonic regression model along the above lines is that usually accurate data on renovation and repair expenditures on a particular dwelling unit between the construction of the initial housing unit and the present period are not available. Without accurate data on repairs and renovations, it will be impossible to obtain accurate estimates of the unknown parameters in the hedonic regression model.

A final practical problem with the above hedonic regression model will be mentioned. Theoretically, "normal" maintenance expenditures could be included in the renovation expenditure terms R_n^t in (116). If this is done, then including normal maintenance expenditures in R_n^t will have the effect of increasing the estimated depreciation rates δ and δ_R. Thus different statistical agencies that have different criteria for deciding where to draw the line between "normal" maintenance and "major" repair and renovations will produce different estimated depreciation rates.

It can be seen that there are many unresolved issues in this area: statistical agency best practice has not yet emerged.

10.5 The Treatment of the Transactions Costs of Home Purchase

Another cost of home ownership needs to be discussed. Normally, when a family purchases a dwelling unit, they have to pay certain fees and costs, which can include:

- The commissions of real estate agents who help the family find the "right" property.
- Various transactions taxes that governments can impose on the sale of the property.
- Various legal fees that might be associated with the transfer of title for the property.

Should the above fees be immediately expensed in the period of purchase or should they simply be regarded as part of the purchase price of the property and hence be depreciated over time in a manner analogous to the treatment of structures in the national accounts?

An argument can be made for either treatment. From the viewpoint of the opportunity cost treatment of purchases of durable goods, the relevant price of the dwelling unit in the periods following the purchase of the property is the after tax and transactions fees value of the property.

[102] Alternatively, if price levels are available for P_S^t and P_R^t from construction price indexes, then these parameters do not have to be estimated.

This viewpoint suggests that the transactions costs of the purchaser should be immediately expensed in the period of purchase. However, from the viewpoint of a landlord who has just purchased a dwelling unit for rental purposes, it would not be sensible to charge the tenant the full cost of these transactions fees in the first month of rent. The landlord would tend to capitalize these costs and recover them gradually over the time period that the landlord expects to own the property. Thus either treatment could be justified and the statistical agency will have to decide which treatment is most convenient from their particular perspective.

11. User Costs for Landlords versus Owners

In the previous section, the various costs associated with home ownership were discussed. Both home owners and landlords face these costs. Thus they will be reflected in market rents and this "fact" must be kept in mind if the imputed rent approach is used to value the services of Owner Occupied Housing. If some or all of these associated costs of OOH are covered elsewhere in the CPI (e.g., home insurance could be separately covered), then the value of imputed rents for OOH must be *reduced* by the amount of these expenditures covered elsewhere.

However, in addition to the costs of home ownership that were covered in the previous section, landlords face a number of *additional costs* compared to the home owner. These additional costs will be reflected in market rents and thus if market rents are used to impute the services provided by the ownership of a dwelling unit, then these extra costs should also be *removed* from the market rents that are used for imputation purposes, since they will not be relevant for owner occupiers. These additional landlord specific costs will be discussed in sections 11.1 to 11.5 below.

11.1 Damage Costs

Tenants do not have the same incentive to take care of a rental property compared to an owned property and so depreciation costs for a rental property are likely to exceed depreciation rates for comparable owned properties. Usually, landlords demand damage deposits but often these deposits are not sufficient to cover the costs of the actual damages that some tenants inflict.

11.2 Nonpayment of Rent and Vacancy Costs

At times, tenants run into financial difficulties and are unable to pay landlords the rent that is owned. Usually, eviction is a long drawn out process and so landlords can lose several months of rent before a nonpaying tenant finally leaves. The landlord also incurs extra costs compared to a homeowner when a rental property remains vacant due to lack of demand.[103] These extra costs will be reflected in market rents but should not be reflected in the user costs of OOH.

11.3 Billing and Maintenance Costs

A (large) landlord may have to rent office space and employ staff to send out monthly bills to tenants and employ staff to respond to requests for maintenance. A homeowner who provides his or her time in order to provide maintenance services[104] provides this time at his or her *after income tax wage rate* which may be lower than the *before income tax wage rate* that a landlord must pay his or her employees. The net effect of these factors leads to higher market rents compared to the corresponding owner occupied user cost.

[103] The demand for rental properties can vary substantially over the business cycle and this can lead to depressed rents or very high rents compared to the user costs of home ownership. Thus imputed rents based on market rents of similar properties can differ substantially from the corresponding user costs of OOH over the business cycle.

[104] Typically, these imputed maintenance costs will not appear in the CPI but if the user cost of an owned dwelling unit is to be comparable with the market rent of a similar property, these imputed labour costs should be included.

11.4 The Opportunity Cost of Capital

The homeowner's *after tax* opportunity cost of capital that appeared in the various user cost formulae considered earlier in this Chapter will typically be *lower* than the landlord's *before tax* opportunity cost of capital. Put another way, the landlord has an extra income tax cost compared to the homeowner. In addition, the landlord may face a higher risk premium for the use of capital due to the risks of damage and nonpayment of rent. However, care must be taken so that these additional landlord costs are not counted twice; i.e., in the present subsection as well as in subsections 11.1 and 11.2 above.

11.5 The Supply of Additional Services for Rental Properties

Often, rental properties will contain some major consumer durables that homeowners have to provide themselves, such as refrigerators, stoves, washing machines, driers and air conditioning units. In addition, landlords may pay for electricity or fuel in some rental apartments. Thus to make the market rental comparable to an owner occupied imputed rent, the market rental should be adjusted downwards to account for the above factors (which will appear elsewhere in the expenditures of owner occupiers).

The factors listed above will tend to make observed market rental prices *higher* than the corresponding user cost for an owner occupier of a property of the same quality. Thus if the imputed rental approach is used to value the services of OOH, then these market based rents should be adjusted downward to account for the above factors.

Although all of the above factors will tend to lead to an *upward* bias if unadjusted market rental rates are used to impute the services of OOH, there is another factor not discussed thus far that could lead to a large *downward* bias. That factor is *rent controls*.

Under normal conditions, the acquisitions approach to the treatment of OOH will give rise to the smallest expenditures, the user cost approach will give rise to the next highest level of expenditures and the use of imputed market rentals will give the largest level of expenditures for owner occupied housing. For the first two approaches, a main driver of the price of OOH is the price of new housing construction. For the user cost approach, another main driver is the price of land. For the imputed rent approach, the main driver of the price of OOH is the rental price index.

The above discussion is far from being complete and definitive but it does illustrate that it is not completely straightforward to impute market rental rates to owner occupied dwelling units. Care must be taken to ensure that the "correct" expenditure weights are constructed.[105]

As can be seen from the material above, the treatment of owner occupied housing presents special difficulties. Astin discussed some of the difficulties that the European Union encountered in trying to find the "best" approach to use in its Harmonized Index of Consumer Prices as follows:

[105] Crone, Nakamura and Voith (2000) have a very useful paper using hedonic techniques to estimate both a rent index and a selling price index for housing in the U.S. They also suggest that *capitalization rates* (i.e., the ratio of the market rent of a housing property to its selling price) can be applied to an index of housing selling prices in order to obtain an imputed rent index for OOH. This is adequate as a first approximation but as the authors note, capitalization rates can change over time (due to changes in nominal interest rates, depreciation rates and expected housing inflation rates). Also, as we have seen in this section and the previous section, actual market housing rents can be expected to be considerably higher than the corresponding imputed rents for owner occupied units of the same quality and hence the use of unadjusted capitalization rates to convert the value of the owner occupied stock of housing into imputed rents can lead to a considerable *weighting bias*.

"A special coverage problem concerns owner-occupied housing. This has always been one of the most difficult sectors to deal with in CPIs.

Strictly, the price of housing should not be included in a CPI because it is classified as capital. On the other hand, the national accounts classifies imputed rents of owner-occupiers as part of consumers' expenditure. This is a reasonable thing to do if the aim is to measure the volume of consumption of the capital resource of housing. But that is not what a CPI is measuring.

Some countries, following the compensation index concept, would prefer to have mortgage interest included in the HICP. This approach could indeed be defended for a compensation index, because there is no doubt that the monthly mortgage payment is an important element in the budget of many households: a rise in the interest rate acts in exactly the same way as a price increase from the point of view of the individual household. But this is not acceptable for a wider inflation index.

So, after many hours of debate, the Working Party came to the conclusion that there were just two options. The first was to simply exclude owner-occupied housing from the HICP. One could at least argue that this was a form of harmonization, although it is worrying that there are such large differences between Member States in the percentages of the population which own or rent their dwellings. Exclusion also falls in line with the international guideline issued 10 years ago by the ILO. Furthermore, it would be possible to supplement the HICP with a separate house price index, which could be used by analysts as part of a battery of inflation indicators.

The second option was to include owner-occupied housing on the basis of acquisition costs, essentially treating them like any other durable. Most secondhand housing would be excluded: in practice the index would include new houses plus a small volume of housing new to the household sector (sales from the company or government sectors to the household sector).

The main problem here is practical: several countries do not have new house price indices and their construction could be difficult and costly. A Task Force is at present examining these matters. Final recommendations are due at the end of 1999." John Astin (1999; 5).

Due to the complexities involved in modeling the treatment of OOH, final recommendations have still not emerged for the HICP.

A fourth approach to the treatment of housing will be studied in the following section. Since this approach has only been applied to owner occupied dwellings, it is not as "universal" as the other 3 approaches.[106]

12. The Payments Approach

A fourth possible approach to the treatment of owner occupied housing, the *payments approach*, is described by Goodhart as follows:

"The second main approach is the payments approach, measuring actual cash outflows, on down payments, mortgage repayments and mortgage interest, or some subset of the above. This approach always, however, includes mortgage interest payments. This, though common, is analytically unsound. First, the procedure is not carried out consistently across purchases. Other goods bought on the basis of credit, e.g., credit card credit, are usually not treated as more expensive on that account (though they have been in New Zealand). Second, the treatment of interest flows is not consistent across persons. If a borrower is worse off in some sense when interest rates rise, then equivalently a lender owning an interest bearing asset is better off; why measure one and not the other? If I sell an interest earning asset, say a money market mutual fund holding, to buy a house, why am I treated differently to someone who borrows on a (variable rate) mortgage? Third, should not the question of the price of any purchase be assessed separately from the issue of how that might be financed? Imports, inventories and all business purchases tend to be purchased in part on

[106] The acquisitions, user cost and rental equivalence approaches can be applied to any consumer durable but of course, to apply the rental equivalence approach, appropriate rental or leasing markets for the durable must exist.

credit. Should we regard imports as more expensive, when the cost of trade credit rises? Money, moreover, is fungible. As we know from calculations of mortgage equity withdrawal, the loan may be secured on the house but used to pay for furniture. When interest rates rise, is the furniture thereby more expensive? Moreover, the actual cash out-payments totally ignore changes in the on going value of the house whether by depreciation, or capital loss/gain, which will often dwarf the cash flow. Despite its problems, such a cash payment approach was used in the United Kingdom until 1994 and still is in Ireland." Charles Goodhart (2001; F350-F351).

Thus the *payments approach* to owner occupied housing is a kind of a *cash flow approach* to the costs of operating an owner occupied dwelling. Possible objections to this approach are that it ignores the opportunity costs of holding the equity in the owner occupied dwelling, it ignores depreciation and it uses nominal interest rates without any offset for inflation. However, if the payments approach is adjusted for these imputed costs, then the result is a rather complicated user cost approach to the treatment of housing. Nevertheless, as was mentioned in Chapter 10, under some conditions, the payments approach to the treatment of owner occupied housing may be a reasonable compromise. In general, the payments approach will tend to lead to much smaller monthly expenditures on owner occupied housing than the other 3 main approaches, except during periods of high inflation, when the nominal mortgage rate term becomes very large without any offsetting item for inflation.[107]

13. Alternative Approaches for Pricing Owner Occupied Housing

For consumer durables that have long useful lives, the usual acquisitions approach will not be adequate for CPI users who desire prices that measure the service flows that consumer durables generate. This is particularly true for owner occupied housing. Hence it will be useful to many users if, in addition to the acquisitions approach, the statistical agency implements a variant of either the rental equivalence approach or the user cost approach for long lived consumer durables and for owner occupied housing in particular. Users can then decide which approach best suits their purposes. Any one of the three main approaches could be chosen as the approach that would be used in the "headline" CPI. The other two approaches could be made available to users as "analytic tables".

We conclude this paper by outlining some of the problems involved in implementing the three main approaches to the measurement of price change for Owner Occupied Housing.

13.1 The Acquisitions Approach

In order to implement the acquisitions approach, a constant quality price index for the sales of new residential housing units will be required.

13.2 The Rental Equivalence Approach

Option 1: Using Home Owner's Estimates of Rents

In this option, homeowners would be surveyed and asked to estimate a rental price for their housing unit. Problems with this approach are:

- Homeowners may not be able to provide very accurate estimates for the rental value of their dwelling unit.

[107] If there is high inflation, then the statistical agency using the payments approach may want to consider adjusting nominal mortgage interest rates for the inflation component as was done in section 10.1 above.

- The statistical agency should make an adjustment to these estimated rents over time in order to take into account the effects of depreciation, which causes the quality of the unit to slowly decline over time (unless this effect is offset by renovation and repair expenditures).[108]
- Care must be taken to determine exactly what extra services are included in the homeowner's estimated rent; i.e., does the rent include insurance, electricity and fuel or the use of various consumer durables in addition to the structure? If so, these extra services should be stripped out of the rent, since they are covered elsewhere in the consumer price index.[109]

Option 2: Using a Hedonic Regression Model of the Rental Market to Impute Rents

In this option, the statistical agency would collect data on rental properties and their characteristics and then use this information to construct a hedonic regression model for the housing rental market.[110] Then this model would be used to impute prices for owner occupied properties. Problems with this approach are:

- It is information intensive; in addition to requiring information on the rents and characteristics of rental properties, information on the characteristics of owner occupied properties would also be required.
- The characteristics of the owner occupied population could be quite different from the characteristics of the rental population. In particular, if the rental market for housing is subject to rent controls, this approach is not recommended.
- Hedonic regression models suffer from a lack of reproducibility in that different researchers will have different characteristics in the model and use different functional forms.
- From the discussion in section 11, it was seen that market rents can be considerably higher than the opportunity costs of home owners and hence using market rents to impute rents for owner occupiers may lead to rents that are too high.[111] On the other hand, if there are rent controls or a temporary glut of rental properties, then market rents could be too low compared to the opportunity costs of home owners.
- There is some evidence that depreciation is somewhat different for rental units compared to owner occupied housing units.[112] If this is so, then the imputation procedure will be somewhat incorrect. However, all studies that estimate depreciation for owner occupied housing suffer from biases due to the inadequate treatment of land and due to the lack of information on repair, renovation and maintenance expenditures over the life of the dwelling unit. Hence, it is not certain that depreciation for rental units is significantly different than that for owner occupied units.

13.3 The User Cost Approach

It is first necessary to decide whether an ex ante or ex post user cost of housing is to be calculated. It seems that the ex ante approach is the more useful one for CPI purposes; these are the prices that

[108] Recall section 8 above.
[109] However, it could be argued that these extra services that might be included in the rent are mainly a weighting issue; i.e., it could be argued that the *trend* in the homeowner's estimated rent would be a reasonably accurate estimate of the trend in the rents after adjusting for the extra services included in the rent.
[110]See Crone, Nakamura and Voith (2000) and Hoffmann and Kurz (2002) for example of such hedonic models that try to cope with the heterogeneity in the rental market.
[111] Again, it could argued that this is a mainly a weighting issue; i.e., it could be argued that the *trend* in market rents would be a reasonably accurate estimate for the trend in home owner's opportunity costs.
[112] "The average depreciation rate for rental property is remarkably constant, ranging from 0.58% to 0.60% over the 25 year period. Depreciation rates for owner occupied units show more variation than the estimated rates for renter occupied units. The average depreciation rate for owner occupied housing ranges from 0.9% in year 1 to 0.28% in year 20." Stephen Malpezzi, Larry Ozanne and Thomas G. Thibodeau (1987; 382).

should appear in economic models of consumer choice. Moreover, the ex post approach will lead to user costs that fluctuate too much to suit the needs of most users. Of course, the problem with the ex ante approach is that it will be difficult to estimate anticipated inflation rates for house prices.

Option 3: The Rent to Value Approach

In this option, the statistical agency collects information on market rents paid for a sample of rental properties but it also collects information on the sales price of these rental properties when they are sold. Using these two pieces of information, the statistical agency can form an estimated *rent to value ratio* for rental properties of various types. It can be seen that this rent to value ratio represents an estimate of all the terms that go into an ex ante user cost formula, except the asset price of the property; i.e., the rent to value ratio for a particular property can be regarded as an estimate of the interest rate less anticipated housing inflation plus the depreciation rate plus the other miscellaneous rates that were discussed in section 10, such as insurance and property tax rates. Under the assumption that these rates remain reasonably constant over the short run, changes in user costs are equal to changes in the price of owner occupied housing. Thus this approach can be implemented if a constant quality price index for the stock value of owner occupied housing can be developed. It may be decided to approximate the comprehensive price index for owner occupied housing by a new housing price index, and if this is done, the approach essentially reduces down to the acquisitions approach, except that the weights will generally be larger using this user cost approach than those obtained using the acquisitions approach.[113] Problems with this approach include:

- It will require considerable amount of resources to construct a constant quality price index for the stock of owner occupied housing units. If a hedonic regression model is used, there are problems associated with the reproducibility of the results.
- Rent to value ratios can change considerably over time. Hence it will be necessary to keep collecting information on rents and selling prices of rental properties on an ongoing basis.
- As was noted in section 11 above, the user cost structure of rental properties can be quite different from the corresponding user cost structure of owner occupied properties. Hence, the use of rent to value ratios can give misleading results.[114]

Option 4: The Simplified User Cost Approach

This approach is similar to that of Option 3 above but instead of using the rent to value ratio to estimate the sum of the various rates in the user cost formula, direct estimates are made of these rates. If the simplified Icelandic user cost approach discussed in section 8 is used, all that is required is a constant quality owner occupied housing price index, an estimated real interest rate and an estimated composite depreciation rate on the structure and land together. Problems with this approach are:

- As was the case with Option 3 above, it will require a considerable amount of resources to construct a constant quality price index for the stock of owner occupied housing units. If a hedonic regression model is used, there are problems associated with the reproducibility of the results.
- It is not known with any degree of certainty what the appropriate real interest rate should be.

[113] Recall the discussion in section 5 above.
[114] However, this is primarily a weighting issue so that the trend in the constant quality stock of owner occupied housing price index should be an adequate approximation to the trend in owner occupied user costs.

- Similarly, it is difficult to determine what the "correct" depreciation rate should be.[115] Moreover, this problem is complicated by the fact that over time, the price of land tends to increase faster than the price of building a residential structure and so the land price component of an owner occupied housing unit will tend to increase in importance which in turn will tend to decrease the composite depreciation rate.

Option5: A National Accounting Approach

This approach makes use of the fact that the national accounts division of the statistical agency will usually collect data on investment in residential housing as well as on repair and renovation expenditures on housing. In addition, many statistical agencies will also construct estimates for the stock of residential dwelling units so that estimates for the structures depreciation rates are available. Finally, if the statistical agency also constructs a national balance sheet, then estimates for the value of residential land will also be available. Thus all of the basic ingredients that are necessary to construct stocks for residential structures and the associated land stocks are available. If in addition, assumptions about the appropriate nominal interest rate and about expected prices for structures and land are made[116], then aggregate user costs of residential structures and residential land can be constructed. The proportion of these stocks that is rented can be deducted and estimates for the user costs and corresponding values for owner occupied residential land and structures can be made. Of course, it would be almost impossible to do all of this on a current basis, but all of the above computations can be done for a base period in order to obtain appropriate weights for owner occupied structures and land. Then, it can be seen that the main drivers for the monthly user costs are the price of a new structure and the price of residential land. Hence if timely monthly indicators for these two prices can be developed, the entire procedure is feasible. Problems with this approach include:

- As was the case with Option 4 above, it will be difficult to determine what the "correct" depreciation rates and real interest rates are.[117]
- It will be difficult to construct a monthly price of residential land index.
- It may be difficult to convert the residential housing investment price deflator from a quarterly to a monthly basis.

All of the above 5 options have their advantages and disadvantages; there does not appear to be clear "winning" option.[118] Thus each statistical agency will have to decide whether they have the resources to implement any of these five options in addition to the usual acquisitions approach to the treatment of owner occupied housing. From the viewpoint of the cost of living approach to the Consumer Price Index, any one of the 5 options would be an adequate approximation to the ideal treatment from the perspective of measuring the flow of consumption services in each period.

References

[115] Due to the lack of information on repairs and renovations, estimated housing depreciation rates vary widely: "One striking feature with the results of all three approaches used in these and related studies is their variability: estimates range from about a half percent per year to two and a half percent." Stephen Malpezzi, Larry Ozanne and Thomas G. Thibodeau (1987; 373-375).

[116] Alternatively, an appropriate real interest rate can be assumed.

[117] However, as usual, it can be argued that errors in estimating these parameters will mainly affect the weights used in the price index.

[118] For consumer durables that do not change in quality over time, Option 5 will probably suffice.

Aizcorbe, A., C. Corrado and M. Doms (2001), "Constructing Price and Quantity Indexes for High Technology Goods", Industrial Output Section, Division of Research and Statistics, Board of Governors of the Federal Reserve System, July.

Astin, J. (1999), "The European Union Harmonized Indices of Consumer Prices (HICP*)", Proceedings of the Ottawa Group Fifth Meeting,* R. Gudnason and Thora Gylfadottir (eds.), held at Statistics Iceland, Reykjavik, Iceland, 25-27 August, 1999, text available at: http://www.statcan.ca/secure/english/ottawa group/ also published in the *Statistical Journal of the United Nations ECE* 16 (1999), 123-135.

Bailey, M.J., R.F. Muth and H.O. Nourse (1963), "A Regression Method for Real Estate Price Construction", *Journal of the American Statistical Association* 58, 933-942.

Beidelman, C. (1973), *Valuation of Used Capital Assets*, Sarasota Florida: American Accounting Association.

Beidelman, C.R. (1976), "Economic Depreciation in a Capital Goods Industry", *National Tax Journal* 29, 379-390.

Böhm-Bawerk, E. V. (1891), *The Positive Theory of Capital*, W. Smart (translator of the original German book published in 1888), New York: G.E. Stechert.

Bureau of Labor Statistics (1983), "Changing the Home Ownership Component of the Consumer Price Index to Rental Equivalence", *CPI Detailed Report*, Washington DC: BLS.

Canning, J.B. (1929), *The Economics of Accountancy*, New York: The Ronald Press Co.

Case, K.E. and R.J. Shiller (1989), "The Efficiency of the Market for Single Family Homes", *The American Economic Review* 79, 125-137.

Christensen, L.R. and D.W. Jorgenson (1969), "The Measurement of U.S. Real Capital Input, 1929-1967," *Review of Income and Wealth* 15(4): 293-320.

Church, A.H. (1901), "The Proper Distribution of Establishment Charges, Part III", *The Engineering Magazine* 21, 904-912.

Court, A. T. (1939), "Hedonic Price Indexes with Automotive Examples", pp. 98-117 in *The Dynamics of Automobile Demand*, New York: General Motors Corporation.

Crone, T.M., L.I. Nakamura and R. Voith (2000), *Journal of Economic and Social Measurement* 26, 153-171.

Diewert, W.E. (1974), "Intertemporal Consumer Theory and the Demand for Durables," *Econometrica* 42, 497-516.

Diewert, W.E. (1976), "Exact and Superlative Index Numbers", *Journal of Econometrics* 4, 114-145.

Diewert, W.E. (1978), "Superlative Index Numbers and Consistency in Aggregation", *Econometrica* 46, 883-900.

Diewert, W.E. (1980), "Aggregation Problems in the Measurement of Capital", pp. 433-528 in *The Measurement of Capital*, Dan Usher (ed.), University of Chicago Press, Chicago.

Diewert, W.E. (2002a), "Harmonized Indexes of Consumer Prices: Their Conceptual Foundations", *Swiss Journal of Economics and Statistics* 138:4, 547-637.

Diewert, W.E. (2002b), "Weighted Country Product Dummy Variable Regressions and Index Number Formulae", Discussion Paper 02-15, Department of Economics, University of British Columbia, Vancouver, Canada, V6T 1Z1. Website: http://www.econ.ubc.ca/diewert/hmpgdie.htm

Diewert, W.E. (2003a), "Hedonic Regressions: A Consumer Theory Approach", pp. 317-348 in *Scanner Data and Price Indexes*, Studies in Income and Wealth Volume 64, R.C. Feenstra and M.D. Shapiro (eds.), Chicago: The University of Chicago Press.

Diewert, W.E. (2003b), "Measuring Capital", NBER Working Paper w9526, Cambridge MA: NBER.

Diewert, W.E. (2003c), "Hedonic Regressions: A Review of Some Unresolved Issues", 7th Ottawa Group Meeting on Price Indices, Paris, May 27-29. Website: http://www.insee.fr/en/av_service/colloques/ottawa/ottawa_papers.htm

Diewert, W.E. and D.A. Lawrence (2000), "Progress in Measuring the Price and Quantity of Capital", pp. 273-326 in *Econometrics Volume 2: Econometrics and the cost of Capital: Essays in Honor of Dale W. Jorgenson*, L. J. Lau (ed.), Cambridge, MA: The MIT Press.

Eurostat, World Bank, International Monetary Fund, Organization for Economic Cooperation and Development, and United Nations (1993), *System of National Accounts 1993*, United Nations, New York.

Eurostat (2001), *Handbook on Price and Volume Measures in National Accounts*, Luxembourg: European Commission.

Garcke, E. and J.M. Fells (1893), *Factory Accounts: Their Principles and Practice*, Fourth Edition, (First Edition 1887), London: Crosby, Lockwood and Son.

Gilman, S. (1939), *Accounting Concepts of Profit*, New York: The Rolland Press Co.

Goodhart, C. (2001), "What Weights should be Given to Asset Prices in the Measurement of Inflation?", *The Economic Journal* 111 (June), F335-F356.

Griliches, Z. (1971a), "Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change", pp. 55-87 in *Price Indexes and Quality Change*, Z. Griliches (ed.), Cambridge MA: Harvard University Press.

Griliches, Z. (1971b), "Introduction: Hedonic Price Indexes Revisited", pp. 3-15 in *Price Indexes and Quality Change*, Z. Griliches (ed.), Cambridge MA: Harvard University Press.

Goldberger, A.A. (1968), The Interpretation and Estimation of Cobb-Douglas Functions, *Econometrica* 35, 464-472.

Gudnason, R. (2003), “How do we Measure Inflation? Some Measurement Problems”, 7th Ottawa Group Meeting on Price Indices, May 27-29, Paris. Website: http://www.insee.fr/en/av_service/colloques/ottawa/ottawa_papers.htm

Hall, R.E. (1971), “The Measurement of Quality Change from Vintage Price Data”, pp. 240-271 in *Price Indexes and Quality Change*, Z. Griliches (ed.), Cambridge MA: Harvard University Press.

Hall, R.E. and D.W. Jorgenson (1967), “Tax Policy and Investment Behavior”, *American Economic Review* 57, 391-414.

Harper, M.J., E.R. Berndt and D.O. Wood (1989), “Rates of Return and Capital Aggregation Using Alternative Rental Prices”, pp. 331-372 in *Technology and Capital Formation*, D. W. Jorgenson and R. Landau (eds.), The MIT Press, Cambridge, MA.

Hicks, J. R. (1946), *Value and Capital*, Second Edition, Oxford: Clarendon Press.

Hoffmann, J. and C. Kurz (2002), “Rent Indices for Housing in West Germany: 1985 to 1998”, Discussion Paper 01/02, Economic Research Centre of the Deutsche Bundesbank, Frankfurt.

Hotelling, H. (1925), “A General Mathematical Theory of Depreciation”, *Journal of the American Statistical Association* 20, 340-353.

Hulten, C.R. (1990), “The Measurement of Capital”, pp. 119-158 in *Fifty Years of Economic Measurement*, E. R. Berndt and J. E. Triplett (eds.), Chicago: the University of Chicago Press.

Hulten, C.R. (1996), “Capital and Wealth in the Revised SNA”, pp. 149-181 in *The New System of National Accounts*, J.W. Kendrick (ed.), New York: Kluwer Academic Publishers.

Hulten, C.R. and F.C. Wykoff (1981a), “The Estimation of Economic Depreciation using Vintage Asset Prices”, *Journal of Econometrics* 15, 367-396.

Hulten, C.R. and F.C. Wykoff (1981b), “The Measurement of Economic Depreciation”, pp. 81-125 in *Depreciation, Inflation and the Taxation of Income from Capital*, C.R. Hulten (ed.), Washington D.C.: The Urban Institute Press.

Hulten, C.R. and F.C. Wykoff (1996), “Issues in the Measurement of Economic Depreciation: Introductory Remarks”, *Economic Inquiry* 34, 10-23.

Jorgenson, D.W. (1989), “Capital as a Factor of Production”, pp. 1-35 in *Technology and Capital Formation*, D.W. Jorgenson and R. Landau (eds.), Cambridge MA: The MIT Press.

Jorgenson, D.W. (1996), “Empirical Studies of Depreciation”, *Economic Inquiry* 34, 24-42.

Katz, A.J. (1983), “Valuing the Services of Consumer Durables”, *The Review of Income and Wealth* 29, 405-427.

Lebow, D.E. and J.B. Rudd (2003), “Measurement Error in the Consumer Price Index: Where do we Stand?”, *Journal of Economic Literature* 41, 159-201.

Leigh, W.A. (1980), "Economic Depreciation of the Residential Housing Stock of the United States, 1950-1970", *The Review of Economics and Statistics* 62, 200-206.

Malpezzi, S., L. Ozanne and T. Thibodeau (1987), "Microeconomic Estimates of Housing Depreciation", *Land Economics* 63, 372-385.

Marshall, A. (1898), *Principles of Economics*, Fourth Edition, London: The Macmillan Co.

Matheson, E. (1910), *The Depreciation of Factories and their Valuation*, Fourth Edition, London: E. & F.N. Spon.

OECD (2001a), *Measuring Productivity: Measurement of Aggregate and Industry-Level Productivity Growth*, Paris: OECD.

OECD (2001b), *Measuring Capital: Measurement of Capital Stocks, Consumption of Fixed Capital and Capital Services*, Paris: OECD.

Palmquist, R.B. (2003), "Property Value Models", Department of Economics, North Carolina State University, Raleigh, NC, June. Forthcoming in *Handbook of Environmental Economics*, Volume 2, K.-G. Mäler and J. Vincent (eds.), Amsterdam: North-Holland.

Reinsdorf, M. (2003), Personal Communication, September 9.

Silver, M.S. and Heravi S. (2001), "Scanner Data and the Measurement of Inflation", *The Economic Journal,* 111 June, F384-F405

Silver, M.S. and S. Heravi (2002), "Why the CPI Matched Models Method May Fail Us", European Central Bank Working Paper 144, Frankfurt.

Silver, M. and S. Heravi (2003), "The Measurement of Quality Adjusted Price Changes", pp. 277-316 in *Scanner Data and Price Indexes*, Studies in Income and Wealth Volume 64, R.C. Feenstra and M.D. Shapiro (eds.), Chicago: The University of Chicago Press.

Solomons, D. (1961), "Economic and Accounting Concepts of Income", *The Accounting Review* 36, 374-383.

Summers, R. (1973), "International Comparisons with Incomplete Data", *Review of Income and Wealth* 29:1, 1-16.

Triplett J.E. (2002), *Handbook on Quality Adjustment of Price Indexes for Information and Communication Technology Products*, OECD Directorate for Science, Technology and Industry, Draft, Paris: OECD.

Walras, L. (1954), *Elements of Pure Economics*, translated by W. Jaffe (first published in 1874), London: George Allen and Unwin.

Wenzlick, R. (1952), "As I See the Fluctuations in the Selling Prices of Single Family Residences", *The Real Estate Analyst* 21 (December 24), 541-548.

Woolford, K. (1999), "Measuring Inflation: A Framework Based on Domestic Final Purchases", pp. 534-543 in *Proceedings of the Measurement of Inflation Conference*, M. Silver and D. Fenwick, Cardiff Business School, Cardiff University, Cardiff CF10 3EU, Wales, UK.

Wyngarden, H. (1927), *An Index of Local Real Estate Prices*, Michigan Business Studies Volume 1, Number 2, Ann Arbor: University of Michigan.

Rósmundur Guðnason[119]: Simple User Cost and Rentals

Abstract

This paper describes and treats the problem of user cost from the point of view of the simple user cost approach to owner occupied housing . The simple user cost method uses a real interest rate as approximation to capital gains and measures depreciation by an inverse geometric rate. The prices are measured by a total house price index.

Usually there is a problem in that connection because of the price divergence between rental markets and changes in house prices although that is not the case in Iceland. The price change in rental markets and house markets do not necessarily have to move in similar fashion. This problem is relevant for every user cost method and the use of the rental equivalence method.

Key words
Consumer price index, cost of living index, household expenditure surveys, owner occupied housing, user cost.

JEL Classification
C43, C81, D11, E31.

1. Market price approach and simple user cost

1.1 Market price approach to owner occupied housing

Using market prices to measure the price change of owner occupied housing is common for two methods that use different approach in estimating the weights share of housing in the CPI. One is based on the flow of services measurement (covering rental equivalence or user cost) and the other is net aqusition where the housing is capitalized as a net cost.

Housing weight for these two methods differs. In the case of the flow of service measurement where rental equivalence is used some countries use information from the National Accounts and other use information collected by the house owner that is asked what rent, in his opinion, he would get for his house, if he rented it out. For simple user cost the annuity (imputed rent) of the average

---

[119] Statistics Iceland

property value, with long term real interest rate and depreciation is used for estimating the expenditure share. Capitalizing the new house is the weight for net aqusition.

Price are measured in these cases by the change in market prices. In the case of rental equivalence it is the market price of rents for comparable housing. In the case of simple user cost it is the market price change of all properties sold independent of whether they are old or new. By the net method it would theoretically be the market price change of new houses that would be used for price measurement. The price measurement is therefore similar and in all cases based on market prices independent of the method of aggregation of weights. It could furthermore be the case that prices for old and new houses moved in similar fashion over time the house price inflation would then be the same for both user cost and net aqusition but if there are differences in weight shares the effect on the total index would differ. Therefore the same index for all properties sold could be used both for net aqusition as for the simple user cost.

According to this market prices are very important element for all methods and the development of reliable price indices of market prices both for rents and houses one of the main issues in the question of owner occupied housing.

1.2 Simple user cost

In the years 2000-2002 about 81 per cent of Icelanders lived in own housing according to the yearly Households Expenditure Survey. Rental equivalence can therefore not be used in Iceland because of the thin and underdeveloped rental market, where the composition of the owner occupied housing stock and the rental market housing is very dissimilar. The only method available to measure the use of the house for living in it is therefore simple user cost.

House is a place to live in and at the same time an investment and to price measure the use is a difficult problem in CPI calculation. The flow of service of living in own house is calculated, as imputed rent in the Icelandic consumer price index, but the buying of the house is an investment and therefore not taken into account directly in the calculation.

Converting a part of investment of a durable (such as a house) into flow of service is done by taking into consideration long term financial (opportunity) cost and the use of the capital (depreciation). The method used is user cost and it combines measure of use of capital and the use of the durable itself, calculated for the duration of year, month or some other time interval. The method to calculate housing as simple user cost[120] and price update the results by the price changes of all properties sold was adopted in November 1992[121]. In the beginning the price measurement was only conducted in the capital area but since March 2000 it was extended to cover the whole country[122]. The base for the calculation is the real estate value of the house (as estimated by the Land Registry) and that information is collected in the yearly Household Expenditure Survey for each household participating.

The user cost is divided into two parts the structure and land but in but is in the model calculated in one figure, with real interest rate that is now approximately 4% and depreciation rate of 1.25% of the house's real estate value. The value of the house is price updated monthly by price index for all properties sold. Owner occupied housing measured in this way covers imputed rent, minor repairs and other cost, such as tariffs for sewerage, garbage and water. The value of the house is collected in the Household Expenditure Survey and is calculated monthly by the simple user cost method.

The simple user cost is calculated in two parts. One is the calculation of weight by using a real interest rate to measure the long term financial cost and the use of the durable. The other part is the price adjustment of the user cost weight (expenditure) by a house price index. Technically it is done by calculating this cost as an annuity[123]. An annuity is a " sequence of equal payments made at equal intervals of time"[124]. In the index calculation the property value is calculated as an annuity and includes both the real interest rate and depreciation. The annuity formula is of the general form:

$$(1) \qquad P_H = A_{HV} * \left[\frac{(1+r)^N - 1}{r * (1+r)^N} \right]$$

where P_H is the present value of the house, A_{HV}, the annuity of the house value, where r is the real interest rate and N the life time of the durable (depreciation is converted to years)[125]. The annuity formula (1) is derived from a geometric series and the interest is calculated over the lifetime of the durable and added to the durables value and then converted into equal payments (annuity). By using

[120] This terminology is due to Diewert (2002) 621 and (2003b) 28 and 53.
[121] Similar user cost method was used in Iceland from 1980 to measure the profitability of the fishing fleet under high inflationary situation.
[122] Correction was made for over estimation of price change of houses in April 2000 lowering the CPI by 0.35 per cent. At the same time under estimation of rent was corrected leading to a 0.34 per cent increase in the CPI.
[123] This user cost method is in some ways similar as Steiner (1961) suggested in the Stiegler report. He uses in his user cost model the annuity method to measure depreciation and interest rates but does not use real interest rates.
[124] Ayres (1963) 80.
[125] The lifetime is calculated as 80 years. The depreciation is 1.25% and there is no scrap value the lifetime is, 1/1.25=80).

annuity both the interest rate and the depreciation is calculated from the same base and increases in the same way by the property index. In addition the rent amount is also calculated over the lifetime of the durable. Lower lifetime of the durable (higher depreciation) leads to lower influence of interest rate changes.

1.2.1 Real interest rate

Nominal interest rates reflect inflation, the higher the inflation, is the higher the interest rates get. This fact that a part of the price of using the capital is due to other factors than the service price for the use of money makes a use of those rates a quality adjustment issue. The quality issue in this case is inflation that is embedded in the interest rates and distorts the interest value making it a quality adjustment necessary. The real interest rate from this point of wiev, the quality adjusted nominal interest rate. The quality adjustment is necessary as is the case of every good and service that has a better or worse quality reflected in its price. The real interest rate can be preset as is the case in Iceland and the inflation measured by adding to a real rate the change in the CPI each month[126].

There is a method available for this kind of correction due to the connection between nominal and real interest rate as expressed by Fisher (1896) identity [127]. The nominal interest rate is designed r_t, the real interest rate as r* and the general inflation rate as p_t. The expression is:

(2) $r_t = (1+r^*)(1+p_t) -1$

It means that the real interest rate, when not known is the difference between the change in the nominal interest rate and consumer inflation and the quality adjustment is expressed by calculating as follows:

(3) $(1+r^*) = r_t/(1+p_t)$

There could be a problem in connection with this in the case of short term movements there are indications that the Fischer effect is not very strong in the short term even if it is so in the longer run[128]. If this is right the use of this method should be extended to some kind of average over a longer period of time.

[126] Indexation of this kindin Iceland refers only to mortages for a period of five years or longer.
[127] Diewert (2003a) 21
[128] Mishkin (1992)

The long term real interest rate used in the simple user cost model calculation shows the return on investment over the lifetime of the durable. In this way the real rate measures the capital gain. It can be lower or higher at periods than the rate of return used but it is approximated by the average long term real interest rate. Methods that measure user cost and do not use a quality adjusted interest rates and in addition take into account all short term movement in house prices will be exaggerating the price changes and not taking into consideration the development of the prices over the lifetime of the durable. The results will therefore also be very volatile[129].

The real interest rates use are nearly fixed but their variable part should reflect in some way the long term trend in the real interest rates and that is why a part of them are kept variable. When setting the average long term real interest rate in the model two main types of financing are viewed. The part that the buyer has to finance by loan and also the required return of his own equity. When consumers buy properties they finance it partly with own equity and partly with mortages. In the simple user cost model the division between this two forms of finance forms are based on information from the sales contracts used for the house price measurement. The own equity rate is fixed in the calculation but the mortgage real interest part is variable. The opportunity financial cost covering the lifetime of the durable is measured in this way.

These shares of finance are used to calculate the real interest rate used in the model. Part of the price of the house is paid by cash and that is approximated as the buyers own equity[130] and is covering more than half of the price given for the house. The real interest rate is of a long term character and is similar to long term real rates used to evaluate the assets of pensions funds. The rate for own equity was therefore set as the estimated rate of return for the pension funds in Iceland and are kept fixed over the durables lifetime. When this calculation method was adopted the long time real interest rate of the pension funds was estimated as 3 per cent and it has been kept unchanged since then[131]. Other long-term real interest rates used are variable over time. These other forms of payment are usually new or old mortgages. The real interest rates used are in the range of 5.0% to nearly 9.0%. These mortages are mainly from the Housing Fond or old loans from the old State Housing Board. Other financing is mainly originating from the pensions funds or the banks.The largest share of these loans have fixed real interest rates that have been unchanged for the period of this method´s lifetime. Hence, the variability of the long term variable real interest rate has been relatively small over time.

[129] Gillingham (1983), Johannessen (2004) and Verbrugge (2004).
[130] This share can partly be financed with loans.
[131] The long time rate of return for the pensions funds is now in the range of 2.0-3.5 per cent. Long time rate of return according to the liability law is 3.5 per cent.

The average real interest rate measured monthly in this way has been around 4% since 1992. The interest rates of the House Fond mortages have bee relatively stable over the last years. In July 2004 this changed and these rates are now set monthly and will probably fluctuate more than before. This led to a change of method going over to a moving five year average for these rates. The motivation was formulated in this way in the press release published on 12 August 2004. ”This practice should ensure that the short term changes in the real interest rates of housing mortages will not cause significant volatility in the monthly measurement of the CPI. On the other hand, it also secures that long term changes in the real interest rates are taken into consideration”. The full text of the press release is found in Appendix 1.

There are three parts that influence the results of the calculation of the annuity. House prices, interest rates and depreciation. The formula for the annuity is:

(4) $$A_{HV} = P_H * \left[\frac{r * (1+r)^N}{(1+r)^N - 1} \right]$$

Dividing through it by $(1+r)^N$ gives the equation in the form:

(5) $$A_{HV} = P_H * \left[\frac{r}{1 - (1+r)^{-N}} \right]$$

According to formula (5) the real interest rate has a direct influence on the annuity and an increase in the average real interest rate, when the rate is 4% and the lifetime is long (80 years) the annuity (imputed rent) increases by almost the same amount as r. The user cost is therefore very vulnerable to real interest rate changes.

The real interest rate has also influence on the value of the property used as the base for calculating the annuity (rent) as lower interest rates lead to a higher house prices. In calculating the present value of the sale contracts the loans with fixed interest rates are discounted by rate of return reflecting the change in the real interest rate. A rise in the real interest rate lowers the present value of the property. This fact is in accordance with the economic reality that a higher real interest rate leads to less demand and lower price of housing.

1.2.2 Depreciation

It is difficult to measure depreciation that should reflect the tear and wear of property and that measurement is always very uncertain. There are generally three methods used to obtain the rate at which structures depreciate. The first method is to observe the age by making a rough estimate of the life of the durable "and then by assuming depreciation model that seems most appropriate"[132]. The second method is to use cross sectional information to set the depreciation rate and the third method is to use information on rental or leasing prices. In deciding the depreciation rate used in the simple user cost calculation the first method was used. "The first and simplest method is to impose a particular depreciation pattern on the average observed life of structures to derive a depreciation rate[133].

The depreciation rate used in the user cost calculation was arrived at mainly by viewing the age of the housing stock. According to the Real Estate Registry the stock at the end of the year 2001[134], divided after the building year, show that 90 per cent of all property is constructed after the year 1940, more than one third in the period 1960-1980 and one third is constructed later. The depreciation rate seems therefore to be in accordance with the property stock divided after age or building year.

The user cost covers both the property structure and the land it is built on. The depreciation for the property is 1.5 per cent setting the life time of the property to approximately 67 years. Land does not wear over time and is therefore not depreciated. The depreciation is calculated of the value of the building. In the price information used to calculate the house price index the price of land is never separated[135]. It is therefore convenient to calculate the depreciation of the whole value of the housing stock, both structure and land. The depreciation rate used in the index is 1.25 per cent of the real estate value.

There are three depreciation methods most common. Straight line depreciation when the depreciation is divided into equal shares, one hoss shay or light bulb depreciation when the durable is depreciated until it falls apart and geometric depreciation when the durables value declines by constant percentage rate. The depreciation is usually in the form $(1-\delta)^N$, where δ is the depreciation rate and N the lifetime of the durable (number of payments). It means that the

[132] Diewert (2003b) 23.
[133] Malpezzi, Ozanne and Thibodeau (1987) 373.
[134] Örn Ingvarsson (2002) 261.
[135] Statistics Iceland has under consideration to calculate separatly user cost from structure and land.

depreciation is largest in the beginning. According to the geometric method the durable is never fully depreciated.

The geometric form in the annuity formula is a inverted geometric depreciation of the type $(1-\delta)^{-N}$. The depreciation calculated in this way differs therefore from the usual geometric depreciation in that it is small in the beginning but increases as the years go on[136]. In addition the durable is fully depreciated which is not the case with the usual geometric depreciation.

It is similar to the one hoss shay method as the depreciation is largest at the end of the durables lifetime and that the durable is fully depreciated but contrary to the one hoss shay method it depreciates over the lifetime of the durable.

2. Rentals and property prices in the consumer price index

Sometimes big increase in house prices are taken as a token that a bubble like behaviour of the market could be at hand which could lead to sudden fall in house prices. One of the things often looked at in that connection is the change in house prices against the price increase of rents. A strong connection can be observed between price changes in the rent market in Iceland and the changes in the market prices of housing, i.e. both indices seem to develop in similar way in the longer run even though they in shorter time periods can move in dissimilar cycles. Imputed and actual rent showed similar changes from March 1997 until November 2003[137]. Because of the small size of the Icelandic rent market it can be said that the house prices lead the rent market.

The rent and property market can move in dissimilar fashion. That could point to imbalances between these two indicators which need not be the case. One reason being regulation or controlling of rent and another could be a quality adjustment issue.

There will also always be difference between the level of market rents and imputed rents. When the rent is set the landlord has to take into consideration the cost incurred[138] such as transaction cost

[136] The depreciation measured as the amortization of the principal, where N = 80, reaches the 50 per cent level in the 64´th year. In the year 73 it covers two third of the total depreciation. The interest payment equals the depreciation amount in the 64´th year and after that the deprication amount is larger than the interest.

[137] The increase in housing prices from March 1997 to November 2003 is nearly 77.9 per cent and the rent index rose in the same period by 78.2 per cent. At the same time the total CPI rose by 28.5 per cent and CPI less housing cost by 22.7 per cent. The CPI is therefore 4.8 per cent higher than the CPI less housing cost in this time interval.

[138] Diewert (2003b) 46-48.

that can be considerable[139]. There is also a quality issue connected to this as owner can adjust their home in any fashion they like unlike tenants that are not allowed to do that.

3. Measurement of property prices

3.1 Real estate value of the property

The Land Registry of Iceland calculates a real estate value for every property in the country. In the middle of the year 2001 the Land Registry revised the estimation method by using hedonic regression after extensive research. The base for the analysis was the capital area and the estimates for other parts of the country were calculated with regional coefficients[140].
The value of all properties in the country are measured in a harmonised way based on information about sold properties. This is done with reference to law as "the law about the measurement of the real estate value says that it should be based on the market price of the property. According the first paragraph. of the law no. 6/2001 the estimated value shall be the discounted market value as estimated last November"[141]. This basic information is the same as used in the price measurement of housing in the CPI and the real estate value is therefore well suited as a base for the user cost calculation.

3.2 Measurement of property prices

Sales contracts are the base for the calculation of the index of property prices. One of the main reasons for using them is the fact that the contracts are standardised for the whole country. Each contract contains information about the property and owners, the sales amount and complete information about the form of payment. Each property has a standardised identity number that is used in the property database that the Land Registry maintains. This very detailed information is the foundation for the valuation of property sales.

The sales information is collected through the Official Registry of Deeds. The formal registry of the change of the ownership is done by the Land Registry and for that the sales contract has to be at

139 Economist (2003) 7.
140 Fasteignamat Ríkisins (2002) 17-22 and Örn Ingvarsson (2002) 259-270.
141 Örn Ingvarsson (2002) 260.

hand. Between 8-10 thousand sales contracts are collected each year covering 8-10% of all properties in the country[142].

The Land Registry of Iceland has collected the sales contracts over a long period of time. The market prices of properties that are gathered from the sales contracts are used by the Land Registry as the base for their evaluation of all house's real estate value. This information is also used in the calculation of the simple user cost in the CPI.

The price concept used in the CPI is cash price. In housing there are different forms of payments used so the price evaluation is more complex. The reason is the fact that there are different forms of payments when a house is bought. As money received today is not the same as money received at another time in the future there is a need to calculate the present value of each contract.

The discount rates varies after the type of payment in accordance with market information. The discounting rate of return is measured monthly and if the change exceed a certain minimum the rate of return is changed. Housing Bonds have the biggest share in housing financing bur will be replaced from July 2004 by ÍLS mortages.

When the discounting rate is lowered, the present value of the property increases. An increase in the discounting rate of return lowers the present value of the property. This is in accordance with the market influence of lower mortgage rates that lead to higher property prices and in the same way as higher mortgage rates lead to lower house prices.

Changes in the market prices and the discounting rate influence the price measurement. The present value of the contract is used for the price updating of properties in the CPI. The price measurement concept is the same as in other parts of the CPI and the prices taken into account are those that the consumer pays in reality for the property. In the long run the nominal and cash house price follow each other but within shorter interval they can part temporarily[143].

3.3 Price index for housing

The house price index is calculated based on the change in the present value of property prices as measured in the sale contracts. The total price information from all the sales contracts are used for

[142] It does not matter weather the number of sales or values are used.

the calculation of imputed rent. The prices used are the average prices for the whole country[144]. In the calculation the combination of the house's size in square meters is kept fixed. The weight is based on the sale's volume in four size categories as it was in the last three years. The calculation is based on three month's moving average with one month's time lag[145]. The sales contracts in April refer to the period January to March and in May for the period February to April e.t.c.. The main indices are four, for houses and apartments inside and outside the capital area. The price change is measured for following type of housing. Houses (13 per cent share) and apartments (59), for the capital area (total 72), houses (15) and apartments (13) outside the capital area (28). The emphasis is on the price change within groups of properties not between types of properties or between regions and the quantity weight between regions is kept constant.

143 From november 2003 to November 2003 house prices in the CPI rose by 11.9 per cent. Nearly half of that change is due to the change in cash prices over nominal prices due to the discounting.

144 This has been in this way since March 2000. The index for the whole country was then calculated back to March 1997.

145 Contracts from places outside the capital area arrive with two months timelag.

Appendix 1

Text of a Statistics Iceland press release, 12 August 2004

Changes in the calculation of owner occupied housing in the CPI in August 2004

The calculation of owner occupied housing in the Icelandic CPI is based on one hand on changes in the market prices of sold properties and on the other hand on the assumed lifetime of the properties (depreciation) and the long time real interest rate (opportunity cost) of the investment in housing

The composition of the real interest rates in the calculation model that Statistics Iceland uses is based on information obtained from sales contracts collected by the Land Registry of Iceland. Approximately two third of the rates do not change from month to month as these are rates on the owners own equity and on old mortages. In the category "other interest rates" the real rates for mortages from the Housing Fund form the largest category and they have been relatively stable over the last ten years.

The lending by the Housing Fund was changed in July 2004 by introduction of loans in cash, ÍLS mortages, that replaced the former housing certificates and bear lower real interest rates. In the calculation of the CPI in July, Statistics Iceland did take into account the lowering of the interest rates of the housing loans as it leads to the lowering of the level of real interest rates. This is in accordance with the practice adopted when similar real rate changes occurred previously, for example at the end of 1993 when the mortage rates were lowered from 6% to 5% and in the beginning of the year 1995 when they rose from 5% to 5.1

In the new lending system, the real interest rates are set each month. For CPI purposes, this could be a challenge to the stability of the real interest rates used in the calculation of the user cost. The long time real interest rates used in the calculation are set in accordance with the lifetime of the house and it would therefore not be appropriate if short term changes in the real interest rates would have major influence on the monthly price measurement. Because of this, it is necessary to extend the model level of real interest rates over a longer period of time.

Statistics Iceland has decided that as of August 2004 the real interest rates used will be calculated as a five year moving average. Initially, the average will be calculated for the period from March 2000 but at that time the house price index was extended to cover the whole country instead of the capital region covered previously. Hence, the first calculation will cover 54 months. The rates will be changed each month and added to the average till they have reached 60 months in February 2005. In each month after that one month will be taken away and a new one added. This practice should ensure that the short term changes in the real interest rates of housing mortages will not cause significant volatility in the monthly measurement of the CPI. On the other hand, it also secures that long term changes in the real interest rates are taken into consideration.

References

Ayres, Frank, JR. (1963), Mathematics of Finance, *Schaum's Outline Series, McGraw Hill Book Company*, May 1963.

Boldsen Hansen, Carsten (2000), The treatment of owner occupied housing in the CPI, final Report by Task Force XII, January 2000, Eurostat.

Central Statistical Office, (1994), Treatment of owner occupiers housing index in the retail price index, *Retail Price Index Advisory Commitee,* December 1994.

Diewert, W. Ervin, (1998), Index Number Issues in the Consumer Price Index, 47-58, *Journal of Economic Perspectives* 12/1,

Diewert, W. Ervin, (1999), The Consumer Price Index and Index Number Purpose, *Journal of Economic and Social Measurement, 27 (2001) 167-248.*

Diewert, W. Ervin, (2001), The Consumer Price Index and Index Number Theory: A Survey, *Discussion Paper 01-02*, Department of Economics, University of British Columbia, Vancouver, Canada, V6T 1Z1.

Diewert, W. Ervin (2002), Harmonized Index of Consumer Prices: Their Conceptual Foundations, *Swiss Journal of Economics and Statistics 138:4, 547-637.*

Diewert, W. Ervin, (2003a), Measuring Capital, *NBER Working Paper w9526*, Cambridge MA:NBER.

Diewert, W. Ervin, (2003b), The Treatment of Owner Occupied Housing and Other Durables in a Cost of Living Index, *Discussion Paper 03-08*, Department of Economics, University of British Columbia, Vancouver, Canada, V6T 1Z1.

Economist, (2003), House of cards, *May 31st 2003, 3-16.*

Fasteignamat Ríkisins (2002), *Árbók 2002*, Fasteignamat Ríkisins 2002 Land Registry of Iceland, Yearbook 2002, (in Icelandic).

Gillingham, Robert (1980), Estimating the user cost of owner occupied housing. *Monthly Labour Review* 1980/February.

Gillingham, Robert (1983), Measuring The Cost Of Shelter For Homeowners: Theoretical And Empirical Considerations. *Review of Economics and Statistics* XLV (2) 1983.

Goodhart, Charles (2001), What Weight Should Be Given To Asset Prices In The Measurement Of Inflation, *The Economic Journal* 111 (June), F335-F356.

Guðnason, Rósmundur (2003a), How do we measure inflation? Some measurement problems, *International Working Group on Price Indicies (Ottawa Group). Proceedings of the Seventh Meeting*, Paris, 27-29 May 2003. Thierry Lacroix editor, 289-320, INSEE, Paris, France, November 2001.

Guðnason, Rósmundur (2003b), Owner Occupied Housing: Market Price Approach To User Cost, Invited paper at the Joint ECE/ILO Meeting on Consumer Price Indices, Geneva, 4-5 December 2003.

The Housing Financing Fund (2003), Process in housing financing, Available at the website: http://ibudalanasjodur.is/displayer.asp?cat_id=15

Hill Peter, (1999), Capital Stocks, Capital Services and Depreciation. Paper presented at the third meeting of the Canberra Group on Capital stock Statistics, Washignton, D.C..

Hulten, Charles R., (1990), the Measurement of Capital, *Fifthy Years of Economic Measurement,* E. R. Berndt and J.E. Triplett (eds.), Chicago: the University of Chicago Press.

Hulten, C. R.and F.C. Wykoff, (1996), Issues in the Measurement of Economic Depreciation: Introductory Remarks, *Economic Inquiry 34,* 10-23.

Johannessen, Randi (2004), Owner-occupied housing in Norway: Why the rental equivalence approach is preferred. Paper presented at the Eight meeting of the International Working Group on Price Indices (The Ottawa Group), Helsinki, Finland, 23-25 August 2004

Jörgenson, Dale W., (1967), the Theory of Investment Behaviour, *in Determinants of Investment Behaviour,* edited by Robert Ferber, National Bureau Of Economic Reasearch, New York.

Jörgenson, Dale W., (1996), Empirical Studies of Depreciation, *Economic Inquiry 34,* 24-42.

Mishkin, Fredric S. (1992), Is the Fisher effect for real? A reexamination of the relationssship between inflation and interst rates., *Journal of Monetary Economics, volume 30, Issue 2, November 1992, pages 192-215.*

Malpezzi S., Ozanne L., Thibodeau T.G., (1987), Microeconomic Estimates of Housing Depreciation, *Land Economics: Nov 1987; 63,372-385.*

Muth, Richard F., (1972), On the Measurement of Shelter Cost for Homeowners in the Consumer Price Index, *Bureau of Labour Statistics, mimeograph, 1972.*

Muth, Richard F., (1975), The User Cost of Owner-Occupied Housing, *Bureau of Labour Statistics, mimeograph, 1975.*

Smith, Dale A., (1975), The Flow of Services Approach to Estimating the Homownership Component of the CPI, *Bureau of Labour Statistics, mimeograph, 1975.*

Steiner, Peter, (1961), Consumer Durables in an Index of Consumer Prices, *Staff Paper no. 6 in The Price Statistics of the Federal Government,.* New York: National Bureau of Economic Reasearch, General series no. 73.

Triplett, Jack. E. (2001), Should the Cost-of-Living Index Provide the Conceptual Framework for a Consumer Price Index?, *The Economic Journal* 111 (June), F311-F334.

Verbrugge, Randall (2004), The puzzling divergence of aggregate rents and user cost, 1978-2001. Paper presented at the SSHRC International Conference on Index Number Theory and the Measurement of Prices and Productivity, Vancouver, Canada; June 30 - July 3, 2004

Örn Ingvarsson (2002), Endurmat Fasteigna í júní 2001, Matsaðferðir við fasteignamat, gerð reiknilíkana og niðurstöður úr þeim, *Árbók VFÍ TFÍ 2002*, 259-270, Reykjavík, (in Icelandic).

Randi Johannessen[146]: Owner-occupied housing in Norway: Why the rental equivalence approach is preferred[*]

Abstract

The Norwegian consumer price index is meant to be a compensation index guided by the economic theory of cost-of-living. The costs of owner-occupied housing are in the Norwegian CPI measured by the rental equivalence approach. The paper focuses why Norway has chosen this approach in spite of the relatively small rental sector in Norway and that the characteristics of the rental households are quite different from the owner-occupied households. Due to the hypotheses that the physical characteristics and the geographic location are similar for both rental and owner-occupied dwellings we believe that the rental equivalence approach provides a good approximation for owner-occupiers. Analyses of the rental sector shows that dwellings owned by private households who typically own one dwelling in addition to their own residence, count for a significant proportion of the rental dwellings. The housing market in Norway has also relatively few professional landlords and almost absence of rent control.

Key words: consumer price index, cost of living index, owner-occupied housing, rental equivalence, user cost function.

1. Introduction

In Norway the cost of living (COL) approach provides a conceptual framework for the Norwegian consumer price index (CPI). An index supposed to provide an approximation to a COL index should ideally relate to the costs of consumption or use, assuming that consumption or use rather than acquisition provides utility. Especially in the case of consumer durables, which can have a lifetime of many years, the distinction between acquisition and use is important. The purchase of a durable can be considered as an investment, designed to provide consumption services over a future time span. For practical reasons, the distinction between acquisition and consumption is ignored in the compilation of the CPI, with one expectation: owner-occupied housing. A house has normally an extremely long lifetime and is according to System of National Accounts regarded as a financial investment.

Buying a house requires very high acquisition costs that usually are spread over time by taking out a mortgage. It is reasonable to assume that living in a house each month generate a flow of consumption services. It is within the COL framework to price the flow of these services - the monthly cost of living in an own dwelling, see for instance Triplett (2001). For rental dwellings the price of the service
- the rent - is covered in the CPI by a monthly rental survey. For owner-occupied dwellings, two common approaches have been analysed:

- Rental equivalence
- User cost function

[146] Statistics Norway

* The author would like to thank Lasse Sandberg for the help during the preparation of this paper.

In rental equivalence approach, one estimates the change in monthly cost for owner-occupied housing by the change in monthly rents for dwellings of similar types that are rented. This principle was introduced in the CPI in august 1967 and is still used with some methodological modifications.

Another way to estimate the cost of owner-occupied housing is the user cost function approach. The user cost approach originates from capital theory that relates the price of houses, the cost of providing housing services and the interest rate for housing investment. The user cost function approach has so far been rejected in the CPI due to the difficulties to implement a reliable method. In this paper some of the difficulties will be emphasized together with some empirical results.

To illustrate the impact of owner-occupied housing based on a user cost function compared to the rental equivalence measure, three alternatives of the user cost approach has been simulated based on Norwegian data; interest rates, house prices, capital gains and depreciating rates. The first version result in quite volatile and even negative user costs due to high capital gains. In this model both unrealised and realized capital gains are included. In the next version, capital gains are reduced to the share of dwellings actually disposed. Thus, only realized capital gains are included in the model. This version leads to less volatile and only positive user costs for the period. The last version is based on expected values of interest rates, house prices and thereby expected capital gains. The only parameter affecting the user costs is the expected price movement of house prices. Thus, the index is really smooth compared with the other versions.

The remainder of this paper is organized as follow. Section 2 and 3 describes the rental and owner-occupied sector in Norway. Section 4 outlines factors that are supposed to determine the rent. In section 5 housing in the CPI is described while section 6 discuss a user cost function as an alternative approach. Section 7 gives some concluding remarks.

2. The rental sector in Norway

After the Second World War, Norway has become a country dominated by owner-occupiers. In 1960 around 40 percent of the Norwegian households were tenants. According to the Population and Housing Census 2001, the rental share is now 23 percent. During the nineties there has been a slight increase of the rental share probably due to rising house prices from 1993.

2.1 The supply of rental dwellings

Norway has a sufficiently active rental sector, though only a minority of Norwegian households are tenants. The stock of rented dwellings is quite heterogeneous as regards both to the distribution of different types of landlords and the size and type of dwellings. Rented dwellings can be divided into the following segments:

- Secondary units in single-family houses or in other house types
- Other private dwellings owned by non-professional landlords
- Public and privately dwellings owned for employees
- Social rental housing
- Private dwellings owned by professional landlords

A secondary unit means that the house is divided into one main and one secondary flat. Many single-family houses in Norway are designed so that a secondary housing unit, normally situated in the first floor or in the basement, can either be let out for hire or used as a part of the resident. Secondary units in single-family house form a quantitatively important part of rental dwellings in Norway. The segment other private dwellings owned by non-professional landlords includes dwellings owned by single persons (or household). Either the owner has an extra dwellings often acquired through inheritance, or he is temporarily absent from his ordinary dwelling.

The two segments "secondary units" and "other private dwellings owned by non-professional landlords" are the dominant types of private rental arrangements in Norway. These have one thing in common; the "landlords" are single persons or households. Furthermore, they typically own only one or a small number of rental dwellings. For owners of secondary units the decision is to letting the house out on hire or use the unit. Letting is correlated to the financial situation, e.g. when the mortgage is high, while own use often corresponds with the family size or when the children are becoming teenagers. Owners of extra dwellings often regard the dwelling as a financial investment and the decision is whether to let out on hire or to sell. Selling often correspond with increasing house prices.

A rental dwelling is classified as being owned by a professional landlord if either a company owns it or if it is privately owned and situated in a multi-family building, e.g. a block of flats, in which the majority of the dwellings are rental units. The share of the rental sector held by professional landlords is quite low in Norway.

Rental dwellings owned by employers (either in private or public sector) and let out for hire to employees also constitute a low share of the Norwegian rental sector. Hospitals and to some extent armed forces mostly own those few dwellings of this kind that are still present in the housing stock. However, this kind of rental dwellings is declining.

Social rental housing also counts for a low share in Norway. According to the Population and Housing Census 2001, rental dwellings where the municipality works as a landlord count for less than 4 percent.

The central idea in this paper is that the stock of rental dwellings in Norway does not consist of a special type of dwelling. It is possible to find both small-occupied units and large rental units and the rented dwellings covers different types of buildings. The geographical distribution of rental dwellings is quite similar in both rural and urban areas. According to Norwegian Social Research (2003), the rental share is 18 percent in the most rural areas while the rental share in the largest cities is just over 25 percent. In smaller towns the rental share various around 21 percent.

2.2 The demand of rental dwellings

The choice of tenure status is believed to be determined as a part of the financial and living situation of a household rather than as characteristics of a dwelling. In Norway it is primarily young people and to some extent the elderly who rent a dwelling. To rent has become a transitional phenomenon for mostly younger people before starting a family and deciding where to live on a more permanent basis. According to the Population and Housing Census 2001, the share of tenants in the age under 25 year is dominating. A relatively high share of tenants is also found among people over 80 years. Around 1/3 of the tenants under age 30 is living with short-term tenancy agreements, while lifelong tenancy is a much more common among elder tenants. During the last 10 years, the share of tenants among younger people has increased slightly, especially in the capital city and other major cities. This is probably due to increasing house prices

3. Owner-occupied sector in Norway

The tax advantage of owner-occupiers is probably one reason why households in Norway do prefer to be owner-occupiers. Owner-occupiers pay income tax based on an imputed rental income. However, this income is calculated in a way that leads to a very low effective tax on the return of owner-occupied housing capital. Before the imputed rental income is calculated a deduction is subtracted from the assessed value. Imputed rental income is set to 2.5 percent of the difference between the assessed value and the deduction of a dwelling. The tax rate for imputed rental income is 28 percent, the same as income from other assets. This rate applies for all taxpayers that are in a position to pay capital income taxes. Interest payment is fully deductible in taxable income and the deductibility is not tied to loans for specific purposes. Owner-occupiers (as well as landlords) can deduct interest payment from the taxable income.

As a main rule, actual income net of operating costs, from rental dwellings is taxed as “normal” capital income with a tax rate at 28 percent. However, owners of second units do not pay tax on the actual income as long as the second unit covers less than 50 percent of the dwelling they actually lives in. In Norway capital gains on rented dwellings are taxed on realization.

According to Statistics Norway's House Price Index, the prices of owner-occupied dwellings[147] have increased by around 70 percent in the period 1989 – 2003[148]. In the early eighties Norway experienced a liberation in the housing market and in the giving of credit. This caused a major increase in the house prices. In the period 1989 - 1992 Norway faced a recession corresponding with a major decrease in the house prices. Since 1993 the house prices has grown by an average of around 9 percent each year. However, after the year 2000, the prices increase more slowly. The development in the house prices is illustrated in figure 1.

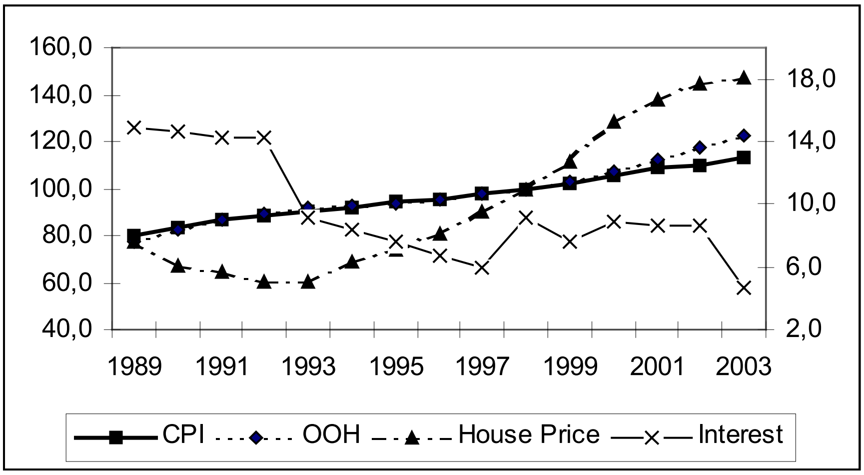
The yearly average nominal interest rate on mortgage accessible from all commercial and savings banks has decreased from almost 15 percent to nearly 5 percent in the period 1989 - 2003, see figure 1.

As illustrated in figure 1, both owner-occupied housing and the overall CPI have had a similar growth from 1989 - 1999. After introducing some methodological changes in the method of estimating the implicit rent of owner-occupiers in 1999, the owner-occupied housing has increased more than the overall CPI.

Figure 1. The development in interest rate, house prices, owner-occupied housing and the overall CPI from 1989 - 2003. Index 1998=100. Interest = change in percent.

[147] Household living in co-operatives are according to the principle of National Accounts, classified as owner-occupiers in CPI. The increase in owner-occupied dwellings includes housing co-operatives.

[148] From 1992 - 2003 the figures are based on Statistics Norway's House price index. From 1989 - 1991 figures from Norwegian Association of Real Estate.



4. Rent determining factors

To a certain extent the rents are a function of the characteristics and geographical location of the dwellings. Other factors that are supposed to determine the market rent of dwellings are:

- Supply and demand of rental dwellings
- The behaviour of the landlords
- The behaviour of potential tenants
- Tenancy agreement (contracts) and rent acts

4.1 Supply and demand of rental dwellings

In a perfect competitive market all agents are small, the products are homogenous and the agents believes that the market price is given and the agents action do not influence the market price. If this was the case in the rental and owner-occupied sector, the dwellings would have been homogenous and consumer world not have preferences over tenure. Both sectors would have been competitive; hence neither the tenants nor the landlords would have power to set the prices.

Much of the rental dwellings in the Norwegian rental sector can be sold to owner-occupiers, and owner-occupied dwellings can be transformed into rental dwellings. Hence, the dwellings are highly heterogeneous yielding thin segments in which the landlords or owners of private dwellings have some sort of market power. The fact that owners of extra dwellings and second units quickly can withdraw from the rental sector will influence the rent of the existing stock of the rental dwellings.

The owners of extra dwellings often regard the dwelling as a financial investment. The decision is whether to let out for hire or to sell the dwelling. The rent must be high enough to cover the operating cost, e.g. maintenance, repairs, insurance, water sewage, chimney sweeping, heating, etc, and give a capital gain at least as high as an alternative investment. With increasing house prices, it is reasonable to assume that the extra dwelling will be sold and thereby reduce the stock of rental dwellings. This could lead to maintaining of the existing rent or even an increase within a specific area. Due to increasing house prices potential buyers of houses may be forced to rent instead of buying, at least temporary until the house prices starts to decrease. Assuming the supply of dwellings fixed, this would lead to increased rents. However, increasing house prices is often corresponded with falling interest rates. If the fall in the interest payment is relatively higher than rise in the house prices, consumers may still afford to buy a house instead of renting.

Professional landlords with many dwellings to let out for hire do not have the same opportunity to immediately sell the dwellings and thereby withdraw from the rental sector. In general, private landlords do not care who the tenants are as long as they are reliable payers. It is assumed that these professional landlords have the highest asking rent for rental dwellings and the most price-determining behaviour. If decreasing demand for rental dwellings, one strategy to maintain the rent at a high level is keeping some rental units vacant for one or more periods.

Owners of single family houses that are split up into a main and a second unit have normally not the opportunity to sell only one of the units. The choice is whether or not to let out for hire the second unit. The owners of second units have an increased tendency of let out for hire when facing high costs, e.g. interest payment or/and operating costs. Keeping all other factors constant, high costs should increase the supply of rental dwellings and thus decrease the average rent level. Simultaneous, if the owners become unemployed, he is likely to rent out for hire to compensate for lower income. Still, it is reasonable to assume that the opportunity to help relatives and/or friends by letting a given part of the house yields a positive utility for the owner. Hence, the rent for second units may be lower than the actual market rent in the specific area.

Recently Norwegian households has experienced repeatedly fall in the interest rate. It is reasonable to assume that this will reduce the supply of rental dwellings like extra dwellings and second units and thus keeping the average rent at the same level or even increase it.

Assuming a tax-reform that keeps the tax-advantage of owner-occupiers but at the same time introducing tax on net rental income[149] from second units at the same rates as net rents from other types of rental dwellings. Full taxation of net rents would most likely decrease the supply of this kind of rental dwellings. Changes in the taxation of the return of housing capital, e.g. a reduction in the tax-advantage of the imputed rental income, will probably increase the demand for rental dwellings.

On very short term, one must assume that both the supply and demand for rental dwellings is given. The market rent will than depend on among others the solvency of potential tenants. A practice of arranging a showing of the dwelling for several potential tenants is common in Norway. The opportunity of viewing the dwelling before renting it is crucial for tenants to decide whether to rent a specific dwelling or not. If a specific dwelling is situated in a very popular area, e.g. near downtown or close to a university, this practice may even cause an increase in the rent because several potential tenants will bid to get the dwelling. However, if a specific rental dwelling is located in a less popular area, this could lead to a negotiating and thus a decrease in the rent.

As mentioned earlier one characteristic of tenants is that renting often is a choice for households who plan to stay for quite a short time in a dwelling. One can assume that short-term tenants are of less concern regarding the actual level as well as the future development of rents

If a tenant wants to change tenure or remains a tenant but wants to move to another rented dwelling, he would face moving costs. It is reasonable to assume that the presence of moving costs causes tenants to move less frequently and that a landlord (or owner of private dwellings) in a situation with decreasing rents thus will maintain the rent level.

4.2 Contracts and rent acts

Contracts in the rental sector are either of fixed duration or by request (at-will). Less than 30 percent of tenants have no contract mostly due to that they rent from family or friends. About one-half of the tenants in the Norwegian rental sector have at-will contracts while around 20 percent have short-term contacts of three years duration or less. The Norwegian Rent Act from 1999 states as a general rule that a dwelling cannot be let out for a shorter period than three years. The tenant can end terminable rental contracts when they want, whereas the landlord faces strict requirement for giving a tenant notice to move out.

[149] Net rental income equals rent minus the operations costs of a rental dwelling.

The act accepts the going market rent for new tenancies. However, according to the act, the market rent is never to be regarded as unfair. Only a rent that is unreasonable high compared to the market rent causes the contract unlawful. Further, the act opens for rent in ongoing tenancies to be adjusted to the market level. This favours the landlord when the rents are rising and the tenant when rents are falling. The Rent Act allows index regulation of the rent on an annual basis based on the CPI.

5. Housing in the CPI

5.1. The rental survey in the CPI

The monthly cost of living in a house is in the CPI covered by a monthly rental survey with about 1,300 responses from tenants. The sample of tenants is drawn from the population of tenants from the Population and Housing Census 2001. The sample is drawn by probability sampling techniques. The sample of rental dwellings corresponded with the tenants in the sample is kept fixed. This means that whenever a tenants move, a new tenant in the specific dwelling should be captured unless the new occupant is an owner-occupier. In this way both new and ongoing tenancies are being captured by the monthly rental survey. A new sample of tenants is drawn annually.

The sample is established through a postal scheme that together with the attributes from the Population and Housing Census gives quite a few characteristics of the rental dwellings as well as some characteristics of the tenants. When the sample is established, the collection of the rents is conducted monthly by a computerized assisting telephone interview. This interview gives an opportunity to ask the tenant several control questions, e.g. about the change in the rent and change in tenure.

5.2. Rental equivalence approach

The rental equivalence approach uses information from the rental sector to estimate owner-occupied housing. It can be argued that the cost of living in an own house cannot be less than the rent that one can receive from a tenant. An owner-occupier always forgoes this amount when he lives in his own house. It can also be argued that over a long period, the cost cannot be greater than the rent of similar rental dwellings, assuming the existence of a sufficiently active rental market, since the owner-occupier always has the possibility of acquiring equivalent housing services at this price.

The rental equivalence approach is incorporated in the CPI given the assumption that the rents charged are a function of the characteristics and geographical location of the dwellings. The approach was adopted in august 1967. However, a major methodological change was conducted when the consumption classification COICOP[150] was introduced in August 1999. Among others, the possibility to identify changes in rents, e.g. publicly subsidized rents that are not representative for owner-occupiers and recognize rental dwellings of similar types to owner-occupied dwellings were improved. The approach is since 1999 used both to derive the expenditure share as well as the implicit rent movement for owner-occupiers. Before 1999, the index of owner-occupiers was assumed to follow the average growth in the rents for both rental- and co-operative dwellings as a whole and the expenditure share was deducted from expenditures on the mortgage interest rate.

There are several strong objections to the rental equivalence approach. Firstly, if rent control and publicly subsidized rents are dominant, changes in rents may not measure monthly costs for owner-occupies dwelling as well. In Norway rent control is limited to a few types of dwellings in limited areas and publicly subsidized rents are not central.

There is also the argument that owner-occupied and rental dwellings are different markets and their prices do not move together. Another strong objection is that rental market is "thin" when it comes to the exact type of housing that is owner-occupied. The primary assumption when introducing the rental equivalence approach is whether data from rental dwellings can be used to develop measures that are representative of owner-occupied units. As discussed above, the Norwegian rental sector do

[150] Classification of individual consumption by purpose.

not consist of a special type of dwelling and that much of the rental dwellings as well as owner-occupied dwellings can easily be converted between tenures.

Analysing the monthly rent survey shows that geographical locations as well as the size of the dwellings gives a significant explanation of the rent level. To derive appropriate expenditures of owner-occupied housing, the data from the rental survey is therefore stratified into geographical regions and the size as well as the type of dwellings.

Based on the rent survey the index of owner-occupied housing in the CPI has increased by around 20 percent for the period August 1999 - December 2003. The index has grown steadily during the period August 1999 to the turn of the year 2002/2003. After this, the index shows a more decreased growth-rate but still positive. January each year shows a greater change than the rest of the months.

6. The user cost approach: an alternative ?

6.1 User cost function

If the rental sector is not representative to provide a good estimate of the cost of owner-occupied housing, the user cost function is an option. One controversial factor is that this approach includes both investment as well as consumption elements.

A simple formulation of the monthly cost for a household, C_m, could be expressed as:

1. $C_m = i_m P_m - [P_m - P_{m-1}] + dP_m$

Rewritten:

2. $C_m = (i_m + d)P_m - \Pi_m P_{m-1}$ where $\Pi_m = \frac{P_m - P_{mt-1}}{P_{mt-1}} = \frac{P_m}{P_{m-1}} - 1$

where i is the appropriate interest rate for housing investment, P is the price of the house itself in month m, Π is equal to the change in the average house price from one month to another, that means the capital gain (or loss) and d is a fixed rate of depreciating - the potential rate of deterioration of the physical capital (the dwelling).

Introducing tax into the user cost function gives the following expression:

3. $C_m = [i_m(1 - \mathrm{T}_t) + d]P_m - \Pi_m P_{m-1}$

Rewritten as:

4. $C_m = a_m P_m - \Pi_m P_{m-1}$ where $a_m = i_m(1 - \mathrm{T}_t) + d$

where T_t is the tax rate in year t.

A rise in the house prices has two effects on user costs. As seen in equation (4) the direct effect occurs because the house price is multiplied by interest and depreciating rate. The second effects is due to that capital gains lower users costs while capital loss will cause higher user costs.

Rearranging expression 4 gives:

If $a_m \frac{P_m}{P_{m-1}} > \Pi_m \Rightarrow C_m > 0$ or $a_m \frac{P_m}{P_{m-1}} < \Pi_m \Rightarrow C_m < 0$

The cost element $a_m \frac{P_m}{P_{m-1}}$ is always positive, while the capital gain (or loss) Π_m is correlated to the development in the prices of dwellings. Decreasing prices of dwellings yields positive user costs, while increasing prices of dwellings cause either a positive or a negative user costs depending on the size of the capital gain. When capital gains are high (typically when house prices are speeding up), they may be large enough to create negative user costs. Once house prices starts to fall, the user costs will become sharply positive. This is one of the main difficulties with the user cost function.

The effect of tax changes depends on the relationship between interest payment and the imputed rental income and the size of these two components. A lower tax-rate will cause lower deducted interest payment. If the interest payment is higher than the imputed rental income, the user costs will increase. However, if the interest payment is lower than the imputed income, the user costs will decrease.

Due to the low effective tax on owner-occupied housing capital, the tax-parameter will be ignored in the simulation of household's user costs. Operating costs such as heating and maintenance are left out as well since they are rather unproblematic to measure. The index from one month to another is thus given by:

$$5.\ I_m = \frac{C_m}{C_{m-1}} = \frac{i_m P_m - \Pi_m P_{m-1} + dP_m}{i_{m-1} P_{m-1} - \Pi_{m-1} p_{m-2} + dP_{m-1}}$$

In the simulation of the user cost below, month will be replaced by quarter since the data used are on a quarterly basis.

The user cost above is a rather simplicity of the real world, assuming active rental and re-sale markets, no uncertainty and no friction cost. When a household is buying a house, normally both a mortgage as well as own equity is required to finance the investment. In the expressions above the mortgage interest rate is assumed to equal the opportunity cost of equity capital[151]. The reason why is the difficulties to estimate an appropriate indicator of the equity capital due to all the existing alternative investments.

An attempt to estimate an actual historic user cost for the period August 1999 to December 2003 has been done. Three different alternatives have been simulated for this period. Version 1 and 2 can be regarded as historical user costs, while version 3 is future user costs based on expectations in July 1999.

In all the three alternatives the market capital value of the stock of owner-occupied dwellings by the end of 1998 as estimated in the National Account is used to get a "reference price level" of owner-occupied dwellings. To get the running market value of the stock of owner-occupied dwellings, each quarter the stock of dwellings is being multiplied with the change in Statistics Norway's House price index.

[151] Gillingham (1983), defines a user cost function where he separates the (unpaid) mortgage and the equity (the amount sum to the price of a house), and introduces an interest rate on the equity as well.

The House price index is used as an indictor of the development in the house prices. This indicator is quarterly and is based on registered purchaser prices of all dwellings sold on the open market during the quarter in question. For co-operatives, both deposit and joint debt is included in the price. To estimate the index, Statistics Norway receives data from the Norwegian Association of Real Estate agents and The Association of Real Estate Undertakings, as well as other real estates agencies that are not members of these associations. The House price index is quality adjusted due to characteristics of the dwellings.

The interest rate is an average of the nominal quarterly mortgage interest rate from all possible commercial and savings banks. This interest rate varies from 4.4 to 9.5 percent during the period. The depreciation rates are in National Account a fixed annual rate of 2.5 percent.

In the first model, the capital gain (or loss) is measured as the change in the value capital of dwellings due to the development in the House price index. That means that both unrealised and realized capital gains are included in the model. Hence the user cost is influenced by the total change in the value of the investments (the dwellings).

It is an important question whether the CPI should be influenced by unrealised value changes. In the second alternative we use the same data as in the first model except that the capital gain is reduced to count only for the share of owner-occupied dwellings that has been disposed in each quarter. This will normally lead to positive user costs. Of course, limit the model to the actual realized capital gains is no guarantee that the user costs always will be positive. If a sufficient part of the stock of dwellings is being disposed, this can lead to quite high capital gains and thereby negative user costs. However, according to the statistics on transfer of properties, the share of the dwellings that are disposed each year is limited to around 10 percent.

A further way of smoothing the user cost is to introduce a long-term fixed interest rate and an expected rate of capital gain. The reason for estimate user costs based on a long-term expected interest rate is due to the assumption that short-term changes in the interest rate are of concern to financing and not housing consumption. Formally such a user cost function can be expressed as:

6. $C_m = \mathrm{E}(a_m P_m) - \mathrm{E}(\Pi_m) P_{m-1}$

where E indicate the expected values of the cost elements, future house prices and capital gains.

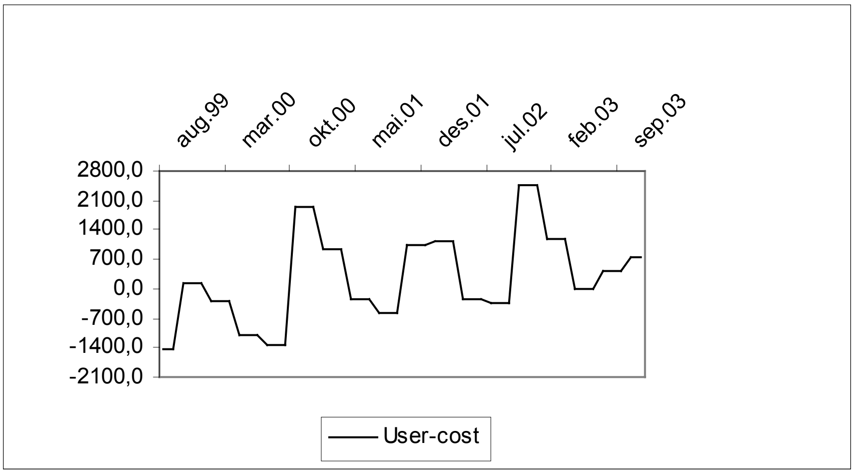
In the version based on expected interest and house prices, none of the data except for the market capital value of the stock of owner-occupied dwellings by the end of 1998 is being used. Given the assumption that owner-occupiers are likely to take a long-term view on their costs, the 10-years interest rate on government bonds is chosen to represent the interest rate of mortgage. The expected trend in the movements of the house prices is based on the historic movement for the period 1989 - 1998. This would normally provide a user cost function that is less volatile and has a more similar trend to the index based on the rental equivalence approach.

6.2 Results

As illustrated in figure 2, 3 and 4, none of the three versions of the user cost track the rent movement that has been incorporated in the CPI in the period August 1999 - December 2003. Appendix 1 gives a briefly overview of the basic findings.

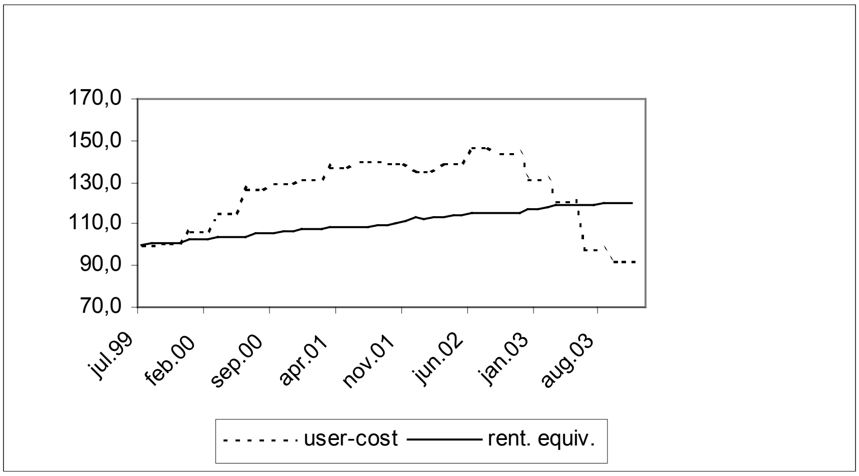
The first version gives (as expected) an extremely volatile index and in some periods even negative user costs due to major increases in the capital gain from one quarter to another. The costs per household are shown in figure 2. The changes in the capital gain are corresponded with the changes in the nominal interest rate. The volatility in the user costs per household is caused by the great variation in the capital gain from one quarter to another.

Figure 2. User cost per household (version 1). Figures in EURO.



By reducing the capital gains to the share of dwellings actual disposed in each quarter, negative user cost is avoided. However, as illustrated in figure 3, the index is still volatile due to changes in the interest rate. The growth in the index is much bigger compared to the rental equivalent approach, except for 2003 where the index is falling. This period correspond with decreasing interest rate.

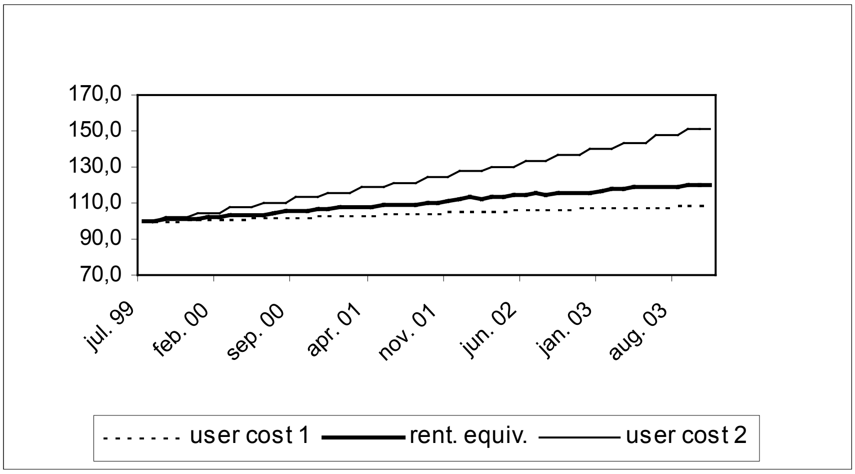
Figure 3. User cost index, version 2. July 1999=100



The third version based on an expected long-term fixed interest rate and expected increase in house prices, is illustrated in figure 4. In July 1999 the long-term interest rate on government bonds was 5,8 percent yearly. The development in house prices and thus the capital gain is based on what household in 1999 would have expected for the next 10 years. However, household expectations are normally based on adaptive expectations, hence the price trends as observed. Assuming that households for the period 1999 - 2009 expect the house prices to develop as they did in the period 1989 - 1999 will give an average increase of only 2.0 percent each year due to the recession in the period 1989 - 1992. If households in 1999 instead did expect the house prices to growth by the same rate as the latest 5-years period (1994 -1998), this would imply an average of almost 10 percent each year.

As seen in the figure, taking the long-term interest rate and the historical average house price trend really smooth the index. Since both the expected interest rate and the expected capital gain are kept fixed, the only difference from one quarter to another is the increase in house prices. If the house prices were expected to move as they did from 1989 - 1998 (2.0 percent on average each year), the third version is contributing less to the overall CPI since the expected house price trend is lower than the actual growth in the rents in the same period. However, if households expected the house prices to growth as they did from 1994 – 1998 (9.8 percent on average each year), the user cost approach would have caused higher monthly costs of owner-occupied housing than the rental equivalence approach. As a comparison the yearly average house prices increase has been around 8 percent in the period 1999 - 2003.

Figure 4. User cost index, version 3. July=100



7. Concluding remarks

The purpose of this paper has been to conduct a theoretical and empirical framework to establish that the rental equivalence approach is appropriate to measure the cost of owner-occupied housing in Norway. The reason why we believe that the rental equivalence is conducting an adequate estimate of the owner-occupied cost is among other that the distribution of rental- and owner-occupied dwellings by characteristics covers essentially the same area in characteristics space.

From an operational point of view, the rental equivalence is also preferable due to the short-term volatility of a basic user cost function. Given the use of the CPI, e.g. escalate both private and public expenditures and judging the impact of observed wage and salary changes upon consumer standard of living, one important question to be answered is whether changes in house prices and interest rates should influence these. From a practical view an index based on less volatile components are more appropriate when it comes to escalate private and public expenditures. Volatile and negative user costs are avoided by incorporating expected future changes in the house prices and a long-term fixed interest rate into the user cost function. However, introducing values based on expectations in the CPI raises the question of whose expectations and how to measure such future changes in an appropriate way.

As discussed by Triplett (2001), the treatment of owner-occupied housing in consumer price indices is a central issue in the debate between advocates and opponents of the COL-theory. For instances Turvey (1999) presents several arguments against inclusion of imputed values to represent the cost of owner-occupied housing. Regarding the Harmonized Consumer Price Index (HICP), Eurostat has rejected the rental equivalence approach because it is considered as imputation. Among others Eurostat argues that imputed rents are the opportunity costs to owner-occupiers of living in their houses rather than a reflection of actual prices faced by them as consumers. However, the HICP is considered as an inflation index and according to Eurostat no opportunity costs are regarded as part of inflation. More information about owner-occupied hosing in the HICP is given in Eurostat (2000).

In the spring 2001, The Central Bank of Norway was given new guidelines for the monetary policy. The central bank is instructed to apply the instruments of monetary policy to establish stable and low inflation. The operational target is an annual growth in consumer prices that over time is near 2.5 per cent. In general, the direct effects on consumer prices resulting from changes in interest rates, taxes, excise duties and extraordinary, temporary disturbances shall not be incorporated into the central bank's monetary policy decisions.

To provide an appropriate target for the monetary policy, Statistics Norway has developed a constant tax index based on the CPI adjusted for both real changes in taxes and energy prices (CPI-ATE). The use of a user cost approach would introduce interest rates into the CPI. Since changes in interest rates should not be incorporated in the monetary policy decisions, owner-occupied housing eventually based on a user cost approach, also must be removed from this constant tax index.

Further work will be done to improve the rental survey within next year. Among others, the sample will be increased and the use of a hedonic rent index will be considered.

References

Eurostat (2000): "Coverage of owner occupied housing in the HICP", CPS 2000/38/6/EN, 38th meeting of the Statistical Programme Committee, Brussels 14 September 2000.

Gillingham, R. (1983): "Measuring the cost of shelter for homeowners: theoretical and empirical considerations". The *Review of Economics and Statistics*, vol. 65(2) (May), pp. 254-65.

Langsether, Å. , Gulbrandsen, L. , Annaniassen, E. (2003): "Leiemarked og leietakers rettsvern". *Norwegian Social Research*, report 2/03.

Triplett, J. E. (2001): "Should Cost-of-Living Index Provide the Conceptual Framework for a Consumer Price Index ?" The *Economic Journal* III (June), F311-F334.

Turvey, R. (1999): "True cost of living Indexes". Paper presented at the Fifth Meeting of the international Working group on Price Indices, Reykjavik, Iceland, August 1999.

Statistical sources:
Norges Bank: "Commercial and savings banks. Average interest rates on housing loan at the end of quarter. Percent per annum". *Economic Bulletin*.
http://www.norges-bank.no/english/publications/economic_bulletin/

Norwegian Association of Real Estate agents,
http://www.nef.no/English.asp

Statistics Norway: "Consumer price index",
http://www.ssb.no/english/subjects/08/02/10/kpi_en/

Statistics Norway: "House price index",

http://www.ssb.no/english/subjects/08/02/30/bpi_en/

Statistics Norway: "Transfer of properties",
http://www.ssb.no/english/subjects/10/14/10/eiendomsoms_e

Eugen Koev: Combining Classification and Hedonic Quality Adjustment in Constructing a House Price Index

Abstract

Statistics Finland has relatively long experience in constructing indices of prices of old flats using both classification and time-dummy hedonic approaches. Each method has proved to have drawbacks. The feasible classification may be too rigid to capture relevant quality changes and the standard time-dummy hedonic approach is not easily interpretable in the context of traditional index number theory. To overcome the disadvantages of these methods the two approaches are combined. A hedonic-method quality adjustment is performed within each cell in a classification and then the index is computed by aggregating cell level quality adjusted prices using an index number formula. It is shown, that each step of the procedure has a very close analogue in the standard practice of statistical offices. A method for evaluating the aggregate contributions of the different characteristics on the overall quality adjustment is developed in order to make the hedonic method more transparent in the context of classical index number theory. Special attention in the discussion is paid to the interpretation of the age profile of house prices, since they are a mixture of two distinct, but empirically non-separable effects having different implication for the quality adjustment.

The method is applied for estimating quarterly indices for Finland during 1987-2000 using very large high-quality register data on all free-market transactions of dwellings in old blocks of flats and terraced houses. It turned out that the quality adjustment of the index was important in evaluation of short-term price movements on thin markets. In the long run the quality adjusted index series do not differ in any substantial way from the simple price averages trend, at least at aggregate levels. The reliability of the index is evaluated by simulation.

The method described in the paper is used in the Finnish official House Price Index 2000=100.

JEL Classification System: C43, E31, R31

Introduction

Most papers related to research in the field of price measurement are, unfortunately, disregarded by official statistics as too complicated and many academic researchers frown at official statistics for their rather unsophisticated nature. This paper is based on past experiences and current research whose results are used in the construction of the quarterly house price index published by Statistics Finland. Hopefully the topic will keep practitioners interested and at the same time the argumentation will satisfy academic readers as well. The discussion of methods used in compiling house price indices is currently particularly relevant, since issues related to prices of owner occupied houses have started to attract attention as it seems likely that housing expenses in owner occupied housing will be included in the European Union's Harmonised Consumer Price Index (HICP) via an index of prices of new dwellings.

As usually in the case of complex goods, the main concern in constructing a house price index is the quality adjustment method and regression methods for quality adjustment have a central place in our discussion. In the context of real estate and house price indices research in the field has been very much influenced by the fact that in the long run more than one sale of a particular dwelling will be observed. The first chapter of the paper is an overview of the mainstream methodology for constructing real estate indices. It covers the repeat-sales regression method suggested by Bailey, Muth and Nourse (1963) and further developed by Case and Shiller (1987) as well as the modern so-called hybrid-type models, for which an already classical reference is Quigley (1995). The chapter contains a discussions of the different assumptions behind these methodologies and points out, that all of them strongly rely on time-invariance of the regression parameters, an assumption which is dubious for long time spans.

The second chapter of the paper is a history of the House price index published by Statistics Finland. The standard classification method and the time-dummy hedonic method used in the past are presented and their shortcomings are discussed from the point of view of official statistics. The author's perception is that a very important concern of statistical agencies is how to link a regression based quality adjustment to the paradigm of classical index number construction methods in a way that ensures good transparency of the results. In chapter three an index construction method based on classification and within-class hedonic quality adjustment is proposed. To enhance the transparency of the procedure a method for evaluating the size of the overall quality adjustment and decomposing it into factors attributable to the different characteristics included in the hedonic regressions is developed.

Chapter four presents the data source, the empirical models and the regression estimation results. A section on the interpretation of the age profile of house prices is included. It turns out to be very important to distinguish between vintage and depreciation effects, which are both captured in the age profile. A method aiming to separate the two proposed by Englund, Quigley and Redfearn (1998) seems to be based on inappropriate identification of the regression coefficients. The author concludes that without outside sample information it is not possible to separate the two effects. It is argued that in the Finnish case interpreting the age profile as reflecting vintage rather than depreciation effect is the less erroneous assumption.

In chapter five the estimated quarterly indices for 1987-2000 are discussed. Data on all transactions of dwellings in housing share corporations provided by the Finnish tax authorities is used. The impact of the quality adjustment procedure is evaluated against the benchmark case in which sample characteristics quality differences are controlled for only by classifying the data. The reliability of the index is evaluated by a simulation method.

It is concluded, that the developed method is applicable generally in cases requiring complicated quality adjustment, such as wage indices.

1.Hedonic Indices for Real Estate Prices: an Overview

Hedonic indices are based on estimating the price relationship between the qualitative characteristics of a complex commodity and its market price. Theoretical foundations of the hedonic method are provided already by Rosen (1974) and Triplett (1983). In the context of housing prices the estimation of hedonic indices and the discussion around them differs in certain aspects from the discussion in the other typical application field of the technique, namely indices for consumer durable goods such as cars and computers. This is due to the differences in the nature of the markets and the data generating processes.

In the case of durable goods the use of hedonic technique is viewed as a means of accounting for the rapid technological quality improvement and the perpetual change of the durable goods on the market. In the case of housing price (or in general real estate) indices the need for regression modelling arise from the nature of the dwellings as goods and the specific features of the housing markets rather than rapid technological change.

House prices react quickly to changes in the economic environment and are considered an important economic indicator, so a real estate price index should be produced at least on a quarterly basis, many users would prefer a monthly one. At the same time the number of real estate sales over a short time period present only a very small fraction of the stock and not even all of them may be available for compiling the index. Each dwelling is endowed with an (almost) unique set of characteristics which determine price, giving rise to very large cross-section price differences. Under these circumstances it is very likely that observed price differences in different time periods as measured by simple statistics such as mean- or median change, or an index based on rough stratification of the dwellings will reflect changes in the quality mix of the dwellings. It is also possible that the quality mix of dwellings is systematically different in economic upturns and downturns, demographic changes may also shift sales towards dwellings endowed with some particular characteristic. One might expect an increasing transaction share in economically vivid regions with growing population. Such phenomena also render price measurement based on simple techniques potentially imprecise and even misleading.

Empirical research on regression methods for property price indices is grossly influenced by the possibility to observe the sales price of the same property (dwelling) at different time periods. In data sets covering a long time span some dwellings are sold more than ones and thus a subset of the data forms a panel. This particular feature of data on property prices has been recognised already in the very early research in the field and efficient utilisation of the panel information has become a major part of the mainstream research programme. In the following we try to summarise the theory.

1.1 The Repeat Sales Model

To introduce the notation, suppose, that the price of a dwelling is determined by the following rather general specification:

$$(1.1) \qquad p_{it} = \alpha_t + \boldsymbol{\beta}_t' \mathbf{x}_{it} + \zeta_{it} + \varepsilon_{it}$$

The subscript i refers to a specific dwelling and t to the time period of the sale. p_{it} is usually the log-price of the property, but it may be some other transformation of the price, $\mathbf{x}_{it}$ are the observed characteristics, α_t is an unknown constant and $\boldsymbol{\beta}_t$ is an unknown parameter vector. ζ_{it} is a dwelling-specific term reflecting idiosyncratic price effects and ε_{it} is a statistical error. Without loss of generality the normalisation $E(\zeta_{it})=0$ can be assumed. Empirical models in the literature are derived by assuming a specific structure of the error term and possibly restricting the parameter vector.

In their pioneering work Bailey, Muth and Nourse (1963) rely completely on a sample of repeat-sales. Suppose that a property is sold in periods τ and t , $\tau > t$. Assume that the property has remained in every respect the same between the two periods, i.e. both its observable and unobservable characteristics have not changed. Assume further that there is no change in the relative market valuation of the characteristics. These assumptions imply that $\mathbf{x}_{i\tau}= \mathbf{x}_{it}$, $\zeta_{i\tau}= \zeta_{it}$ and $\boldsymbol{\beta}_\tau = \boldsymbol{\beta}_t$. In this case the price difference of the dwelling i between τ and t is:

$$(1.2) \qquad p_{i\tau} - p_{it} = (\alpha_\tau - \alpha_t) + (\varepsilon_{i\tau} - \varepsilon_{it})$$

This extremely simple model can be estimated by ordinary least squares (OLS) under the assumption that $(\varepsilon_{it} - \varepsilon_{it})$ are independent with zero mean and constant variance. By appropriately normalising the coefficient for the initial period α_0, one gets the desired index directly from the coefficients α_t.

A more sophisticated variation of (1.2) proposed by Case and Shiller (1989) has become a benchmark case in the literature. Case and Shiller argue, that the property-specific value captured in the term ζ_{it} in equation (1.1) experiences random shocks over time:

(1.3) $$\zeta_{it} = \zeta_{i(t-1)} + v_{it} = \zeta_i + \sum_{j=1}^{t} v_{ij}$$

where ν_{it} is white noise and hence ζ_{it} is a random walk. Under these assumptions (1.2) becomes

(1.2') $$p_{i\tau} - p_{it} = (\alpha_\tau - \alpha_t) + (\nu_{i\tau} - \nu_{it}) + (\varepsilon_{i\tau} - \varepsilon_{it})$$

In the specification (1.2') the error term has two components and is heteroscedastic, since the variance of $(\nu_{i\tau} - \nu_{it})$ depends on the length of the interval (τ-t). Case and Shiller develop an obvious feasible generalised least squares (GLS) estimator, in which the squared OLS residuals of (1.2') are used in an auxiliary regression on a constant term and the interval (τ-t) to obtain estimates of the variances of ν_{it} and ε_{it}. These are then used to form the appropriate weight-matrix and re-estimate (1.2') by GLS.

The appeal of the repeat-sales model is that there is no need to know anything about the price-characteristics relationship, such as what are the relevant characteristics and the correct functional form. In this sense specifications (1.2) and (1.2') are robust. However, it assumes that the price-characteristic relation, whatever it is, is constant over time and that all characteristics of the properties are unchanged[152]. A very serious problem with the procedure is that it is extremely wasteful in terms of observations, since price information for properties whose selling price is observed only once is simply thrown away. An even more serious objection to the method is that properties sold more than once within a short time period are likely to be a non-random and non-representative subsample of all sales and hence the estimated index will be biased. These drawbacks of the method have been recognised in the 90's and have lead to much empirical work, whose purpose has been to develop procedures, which recognise that a subset of the data forms a panel (multiple sales of the same property), but makes use of the information of properties sold only once. Such methods are normally referred to in the literature as "hybrid".

1.2 The "Hybrid Models"

The insight of the "hybrid" models is that the panel of properties sold more than once may provide, under certain assumptions, the possibility for more efficient estimation than simply a pooled OLS procedure will do. Assume that $\boldsymbol{\beta}_t$ is time invariable, i.e. $\boldsymbol{\beta}_t = \boldsymbol{\beta}$. Then equation (1.1) becomes

[152] In research based on "pure" repeat-sales methods this assumption is normally checked and only observations satisfying it are selected from the data

$$(1.4) \qquad p_{it} = \alpha_t + \beta' \mathbf{x}_{it} + \zeta_{it} + \varepsilon_{\text{it}}$$

Under standard assumptions for the structure of the composite error term $\zeta_{\text{it}} + \varepsilon_{\text{it}}$, pooling all observations and using OLS will provide consistent estimates of the regression parameters. However by recognising the panel structure of the repeat-sales subset of the data, estimation can be improved in certain cases. The seminal paper of Quigley (1995) shows how the repeat-sales can be used first, to distinguish between effects of observed characteristics in the vector $\mathbf{x}_{it}$ and unobserved ones captured in the idiosyncratic term ζ_{it}, and second, to improve estimation efficiency by appropriate GLS procedure. Following Case and Shiller, Quigley (1995) assumes that the idiosyncratic term is described by (1.3), that is $\zeta_{it} = \zeta_i + \sum_{j=1}^{t} v_{ij}$ where ν_{ij} is essentially a white noise[153]. Then by rearranging the error terms of (1.4) and defining $\eta_{it} = \sum_{j=1}^{t} v_{ij} + \varepsilon_{it}$ one obtains

$$(1.4') \qquad p_{it} = \alpha_t + \beta' \mathbf{x}_{it} + \zeta_i + \eta_{it}$$

The following structural assumptions are made: $E(\zeta_i)=0$, $E(\eta_{it})=0$, $E(\eta_{it}\,\eta_{j\tau})=0$, $E(\zeta_i\zeta_j)=0$, $E(\zeta_i\eta_{jt})=0$, $E(\zeta_i)^2=\sigma^2_\zeta$, $E(\eta_{it})^2=\sigma^2_\eta$ for all $i\neq j$ and $t\neq\tau$ and $E(\eta_{it}-\eta_{i\tau})^2$ is a quadratic polynomial of $(t-\tau)$. The proposed estimation procedure is the following. First (1.4') is estimated using the sample of repeat-sales. The idiosyncratic effects are explicitly modelled by including property-specific dummies. The squared residuals of this model are then used to estimate σ^2_η and the time structure of $E(\eta_{it}-\eta_{i\tau})^2$. Then the model is estimated using again the same repeat-sales subset, but this time without the property-specific dummies. The squared residuals of this model together with the estimates of σ^2_η and the time-structure of $E(\varepsilon_{it}-\varepsilon_{i\tau})^2$ from the previous step are used to estimate σ^2_ζ. The structure of the variance-covariance matrix of (1.4') is now completely identified and in the final step all observations are used to estimate the equation by GLS.

1.3 Discussion

Different estimators are consistent and efficient under different conditions and choice of the most appropriate one is largely dependent on what particular problems the researchers think are of primary importance for the phenomena under study in general and for the particular data set used. Repeat-sales methods avoid a lot of modelling uncertainty related to omission of relevant explanatory variables and functional specification of the regression. Although bias due to non-randomness of the repeat-sales is recognised as a possibility by Quigley (1995), the hybrid method is primarily concerned with efficient use of all sample information and not with sample selection bias. This is because if the subset of repeat-sales in a data is a result of a non-random process, then the estimation results of the first step of Quigley's procedure, which relies on repeat-sales, are potentially biased and it is not self-evident whether the overall estimation is consistent. However, efficiency is achieved by imposing explicit structure of the regression function, a feature which pure repeat-sales methods avoid.

There are some other aspects of the hybrid model worth discussing. The procedure can be interpreted as an estimation of an unbalanced random-effect panel model. The central assumption of such models is that the two error terms ζ_i and η_{jt} are both non-correlated with the observable variables The efficiency gain of the hybrid model estimator over either pooled regression or repeat-sales models comes from utilisation of both between-unit and within-unit price variability. Pooled regression utilises only between-group and repeat-sales only within-group variability (see Greene,1997, 618-620). Knowing this fact, it is not surprising that a normal empirical finding is that the hedonic index based on hybrid models has narrower confidence intervals as compared to estimates obtained from either pooled regression or repeat-sales.

Estimation of both the repeat-sales and the Quigley's hybrid model heavily relies on the assumption that the relative valuation of the different characteristics of the property are constant over time, implying time-constant parameter $\boldsymbol{\beta}$. This assumption is very restrictive for analysing data collected over 10-years long time period as is the case e.g. in Case and Shiller (1979), Quigley (1995), Englund, Quigley and Redfearn (1998 and 1999). Meese and Wallace (1997) test different assumptions of the repeat-sales and the hybrid model. One conclusion is that estimates of β based on either repeat-sales or single sales are statistically different. The result is interpreted by the authors as supporting the hypothesis that repeat-sales prices are unrepresentative. The hypothesis of time-constant β is also rejected in this research based on data from two cities in 1970-1988. The authors suggest as an alternative a non-parametric regression specification allowing for time-variable parameters.

[153] I slightly re-interpret Quigley's (1995) presentation and don't go into full detail, he assumes for example a more complicated structure of the variance of the term ν_t, but this is not relevant for the discussion.

As obvious, there is not necessarily a best solution, since under different assumptions different procedure may be best. The author's opinion is that arguments against use of pure repeat-sales models are compelling. For time-periods of several years the assumption of parameter constancy in the case of houses is reasonable, so a hybrid-model of the type that Quigley (1995) suggests might be considered a good solution for short index series. However, the efficiency gain of utilising the repeat-sales information in estimation of the index is proportional to the fraction of repeat-sales in the overall data. The shorter the time span over which estimation is performed, the smaller the repeat-sales fraction and the smaller the benefit. Moreover, repeat-sales over short time periods are most likely subject to severe selection biases (in the very long run any house will be eventually sold more than once). For long time periods the hypothesis of parameter constancy has no a priori grounds, and as the work of Meese and Wallace (1997) shows, is likely to be violated. This considerations suggest that research in the field should perhaps be focused more on flexibility and robustness of the proposed methods rather than on their relative efficiency under perhaps unduly stringent assumptions.

From the point of view of official statistical agencies, there are several aspects to be considered when interpreting the research results and their suggestions. This requires some clarification of the language and the normal practice of statistical offices. The starting point of classical price index theory is that there exists a complete list of all goods and their unit prices. The items in that list are the same at all times. To make the concept operational, statistical offices would devise a classification of the goods into homogenous classes, so that goods within each class are (presumably) close substitutes. Today the classifications are increasingly co-ordinated at international level. Then from within each class one or several precisely specified items will be selected and their unit prices will be followed e.g. each month. The choice is normally made on some notion of representativeness of the items such as market share in the class. The observed prices of an item in the base and comparison periods form a “matched pair”. The precise specification of the items to be followed ensures, that the matched pair unit-price ratio reflects only real price changes and not quality differences. If more than one item from a class is selected, then the price change estimate for the class is obtained by simply averaging the matched pair price ratios, usually without weighting. The price index is then computed by aggregating class-level price- and quantity information using an index-number formula, which is almost always the Laspeyres.

There are two distinct parts in this basic process for computing official price indices. First, there is the quality control part, achieved through the classification and the items selection from each class. This ensures that the index is comprehensive and that observed unit price changes are legitimate measures of price changes. At that level there is no index number problems involved, at least from the point of view of traditional index number theory. The second part is the aggregation. This is the focus of classical index theory, which defines and examines the properties of index formulae. A typical side-process is the calculation of a number of sub-aggregate level indices for different sub-sets of interest (i.e. food-price index in the consumer price index (CPI)) alongside with the "total" price index. A matter of interest and concern for the statistical agencies is the coherence between the total price index and the sub-indices and it requires, among other things, careful planing of the classification to ensure that a sub-index coverage is an union of distinct classes.

The need for quality adjustment in official indices arises when the above discussed basic quality-control procedure fails. In the CPI this is nowadays typically the case of household appliances and especially computers, since models on the market disappear very fast and are replaced by more sophisticated new ones. A quality adjustment procedure, as understood in statistical offices, should tackle that type of problems and should not affect other aspects of the index compilation such as aggregation issues.

An official index for house prices must be interpretable within the established paradigm, which, after all, has sound foundations in classical index number theory. From this point of view the real estate price indices described in the literature should be viewed as estimation techniques aimed at obtaining quality adjusted price ratios within some reasonably homogenous class of houses. Researchers typically provide computations for well defined geographical regions within which different houses can be viewed as reasonable substitutes. However the treatment of the topic is usually heavily concentrated around estimation techniques and econometric issues, the links to standard index construction practices and issues such as computation of an overall index are by-passed as self-evident or outside the scope of the research.

On the other hand, some recent research topics such as the so-called time-aggregation bias (Calhoun, Chinloy and Megbolugbe (1995), Englund, Quigley and Redfearn (1999)) , are not relevant from the standpoint of classical index number theory. The problem in a nut-shell is that prices of houses may change significantly from e.g. month to month and a quarterly index will fail to report such changes. Such smoothing of the series cannot be named a bias in the paradigm of classical index theory, where the length of the base and the comparison period can be freely determined. Of course, the loss of information on within-quarter price changes is highly relevant e.g. for purposes of investment risk analysis as pointed out in Calhoun, Chinloy and Megbolugbe (1995), but it nevertheless does not relate to the notion of bias as normally understood by index number experts.

Most important, the existing research is not explicit enough about which characteristics and how affect the quality adjustment. Although theoretical foundations of the hedonic method are provided already by Rosen (1974) and Triplett (1983), only very recently Triplett (2001) and Diewert (2001) have shown, for example, that the matched pairs method normally used in statistical offices can actually be expressed in a hedonic regression form. This seems to be a major reason why statistical agencies are still not comfortable with quality adjustment based on regression methods. Because the author's belief is that this line of research is crucial for the wider acceptance of hedonic quality adjustment in official statistics, an important part of the discussion in this paper is focused on how tangible estimates of the size and the magnitude of the quality adjustment of the characteristic included in the model can be obtained.

2. The History of the Finnish Official House Price Index

2.1 Main Features of the Finnish House Market

Around three fourths of the net wealth of Finnish households is in housing, meaning that home ownership is the only considerable wealth asset for most Finns. The proportion of owner-occupancy rate in Finland has declined in the 90's but is still as high as 60 percent. The decline in owner-occupancy rate is explained largely by the abolishment of rent controls and to a smaller extent by the appearance of new forms of occupancy, which stay between ownership and renting.

Single family houses and apartments in blocks of flats form 40 percent of the housing stock each, apartments in terraced houses account for 13 percent of the stock and the rest of the dwellings are apartments in other than residential buildings. The volume of new housing has been rather low since the end of the 80's, at present new dwellings form about one percent of the stock per year and at least 1/3 of them are intended for rental use. Building companies are involved mainly in the construction of blocks of flats and terraced houses, while most single family houses in Finland are still built using mainly family's own labour.

An institutional peculiarity on the Finnish housing markets is the rather technical distinction between real estate ownership and ownership of a dwelling in a housing share corporation. The land and the buildings in the case of blocks of flats and terraced houses are owned by so-called housing share corporations (asunto-osakeyhtiö). The owner of a dwelling formally owns those shares in the company, which explicitly entitle her rights to a specified dwelling. Single family homes typically form real estates, but for taxation reasons some single family houses are also organised as housing share corporations.

House markets transactions are concentrated in the bigger cities and apartments in blocks of flats account for over 50 % of the transactions of old (existing) dwellings. Single family house is transacted in only about 10-20 % of the cases. Most buyers buy the dwelling for own use.

Finnish house markets are believed to be very sensitive to interest rates, since a purchase of a dwelling is often financed by a loan, whose interest rate is linked to some short term market rate such as the six or twelve month Euribor. House loans are issued by deposit banks and in the past had short repayment period, five to ten years. The situation changed in the 90's and currently a loan can be normally taken with repayment period of 25 and even 30 years.

The Finnish house markets have been very volatile in the past 15 years. During the second half of the 80's liberalisation of the loan market together with booming economy lead to extremely strong price upsurge. In the subsequent recession prices fell to about half of their 1988-1989 levels. As the economy recovered, prices started to grow fast again in 1996. Rising housing prices were supported by interest rate stability secured by Finland's EU-membership and the government's commitment to the monetary union. A strong internal migration towards the capital region lead to scarcity and above-the-average housing price increases in Helsinki and the surrounding municipalities. Towards the end of 2000 the long upward trend in house prices at least temporarily bent down.

2.2 Data Sources and Definitions of Statistics Finland's House Price Index

The importance of the house market developments is widely recognised in Finland and house prices are an important topic of interest for politicians, social planners, economy analysts and ordinary people as well. Against this background it seems natural that Statistics Finland publishes a quarterly house price index since 1985. Afterwards computed historical series are available for some regions from 1970 onwards. Up to the end of 2001 the index was based on actual transaction prices of dwellings in housing share corporations, in practice apartments in blocks of flats and terraced houses. Information was gathered from the major real estate agents and covered about 1/3 of all transactions in that type of housing, but regionally and at different times the coverage has varied from 20 to 70 percent of all transactions. The data source and the methodology were renewed at the beginning of 2002. The new data source and the method are described in the following sections of the paper.

The old and the new index are based on the same definitions. The index covers transactions of dwellings in housing share corporations. Detached family houses forming a separate real estate are not included. The reason is twofold. First, real estate transaction information is registered separately. Second, since only a small number of real estate are carried out quarterly it will not be possible to reports the index at the same regional breakdown as the current one. Statistics Finland has started to produce a separate index for single family houses at rough regional level since the beginning of 2002. Over the last 15 years the price movements in single family houses have followed extremely closely the price movements of the dwellings covered by the house price index.

The house price index excludes new dwellings again because only few transactions are reported quarterly. Statistics Finland is working on the possibilities to collect more representative data on new house sales. Since there is no information on whether a flat is sold by the building company to a private person/household, all flats whose year of completion is the current or the previous calendar year are considered new and the ones with earlier year of completion are defined as old.

At present the house price index contains a series for the whole country (Åland not included), the provinces, major towns and other interesting geographical entities such as Greater Helsinki, comprising the municipalities of Espoo, Helsinki, Vantaa and Kauniainen. Within the regional breakdown separate series are published for type of building (block of flat or terraced house) as well as by number of rooms for the apartments in blocks of flats. The regions for which separate series are currently computed is presented in the Appendix. All results presented in the paper follow the definitions of the official index.

2.3 The Pure Classification Index

From 1985 till the second half of 1995 the house price index was compiled using a pure classification approach. Data was stratified first by region and within the region in the following way:

	Apartments in blocks of flat			Terraced house
Year of completion	1 room	2 rooms	at least 3 rooms	
up to 1960				
1961-1970				
after 1970				

Thus a classification by (region) X (type of building)X(number of rooms)X(age group) was created. Each cell in this classification was treated, in the terminology of the Harmonised Index of Consumer Prices (HICP), as an elementary aggregate. Simple arithmetic price ratios between the base and the comparison period were compiled and then aggregated into an index using the Laspeyres formula. The weights[154] were determined by the number houses in each elementary aggregate in 1980.

If in any particular period the number of transactions in an elementary aggregate cell was 0, then the index was computed under the assumption that the prices between the current and the previous quarter in the cell had remained unchanged. We shall argue later on, that this procedure does not bias the index unless some class with positive weight is constantly empty.

After several years of experience with this method it was considered, that it is not necessarily adequate for the smaller regions since the index series there exhibited too much quarter-to-quarter fluctuations possibly caused by uncontrolled quality variation of the flats sold. Since no further classification of the data was feasible, the solution was seen in using hedonic methods, which potentially allow to control for more quality characteristics than classification at the cost of imposing more structure on the way these enter the house price formation (i.e. the regression functional form). Thus the method used until the end of 2001 and described in the next section was adopted.

2.4 Time-Dummy Hedonic Regression

[154] Strictly speaking, the Laspeyres weights should be value shares

The variation observed in the price per square metre is grouped into three main factors: differences in the individual characteristics of the dwellings, the effect of location and residential area, and the sales period. The model applied is of the form:

$$(2.1) \quad Ln(p_{ati}) = \beta_0 + \sum_{k=1}^{K} \beta_k x_{kti} + \sum_{a=2}^{N} \mu_a A_{ati} + \sum_{t=2}^{J} \lambda_t T_{ti} + \varepsilon_{ati}$$

The explanatory variable is the logarithmic price per square metre $ln(p_{ati})$, where sub-index i refers to the observation, a to the location and t to the time period. x_{kti} refers to the value of characteristic x_k in observation i at period t. A_{ati}'s are the location dummy variables which receive the value 1 if dwelling i is located in area a. T_{ti} ´s are the time dummy variables,.β_0 β_1 ...β_K, μ_2...μ_N, λ_1 ... λ_J are the regression parameters.

Estimation was based on data for eight quarters: the seven most recent and a fixed "base-period" quarter. The estimate of the price trend was obtained from the coefficients of the time dummy variables in the model. Since the explanatory variable was logarithmic, the coefficient estimate of the dummy variable indicated the change in price level in comparison with the "base" period in log-percents. The index point number was computed via the formula $I^t_0 = I_b e^{(\lambda_t)}$, where I^t_0 denotes the index point number for period t, I_b is the index point number in the base quarter and λ_t is the estimate of the coefficient of the time dummy variable for period t.

The locations in the model were identified at postal code area level. Other explanatory variables were floor area of the dwelling and its square root, number of rooms, age of the dwelling and its square root, type of building and physical condition indicators.

Equation (2.1) is a very standard specification in the hedonic index literature going back to Griliches (1971) and still very popular. A particular feature of Statistics Finland's implementation was that instead of weighting results from separate regional regressions to obtain aggregate level indices, equation (2.1) was independently estimated for all subsets of the data for which an index was to be compiled. This feature lead to incoherent series. The problem is the following. If say, in Helsinki, prices for apartments in blocks of flats rose by 5 % from the previous quarter and prices of apartments in terraced houses rose by 3 %, then one would expect that the average price rise for all types of flats will be between 3 and 5 percent. This is not guaranteed by the described method, since the regression coefficients, in particular the coefficients of the time dummies for blocks of flats, flats in terraced houses and all flats are estimated by running separate, independent regressions.

It should be noted that this type of non-coherency problems are not uncommon and some famous classic index formulas such as the Fisher formula can certainly exhibit the same problem. A more subtle but similar problem is related to the fact that in a chained index the overall index change between two periods may be smaller or greater than the minimum or maximum change of its sub-indices. In practice, such situation may not exist or may not be noticed in official price statistics, especially in the CPI, first because CPI uses the Laspeyres formula, which is consistent in aggregation, second, the CPI is usually divided into at least 5-10 sub-indices and third it is seldom chained.

Obviously, appropriate weighting solves the coherency problem. However it will not remove the low transparency of the procedure. If one wishes to compare the results of the hedonic equation of the form (2.1) with say, an index based on classification etc., one would be able to state in what way the series differ and may suggest some plausible explanations, but no tangible numeric evidence of what regressor and how affected the quality adjustment and the index can be provided. The situation is to be compared with the normal practice of the CPI, where, say, the evaluation of the impact of gasoline price changes on the overall index is a standard procedure. So one has difficulties to demonstrate that the regression does its job appropriately especially in situations where the hedonic index persistently moves faster or slower than some unadjusted measure such as the simple price average. It makes it then rather understandable, that statistical offices feel uneasy about hedonic solutions of the above type, because the credibility of the statistics requires that there is some accounting of what was adjusted and why. The method proposed in the following section addresses this issue.

3. Methodology of the New Index

In view of the preceding discussion it is natural to search for combined methods that will eventually retain the good features of both the classification and the hedonic approach but will mitigate the problems associated with either of them. What is suggested here is first classifying the data with respect to the characteristics along which most price variation is observed. Then regression analysis is used to do cell-specific quality adjustment with respect to other important characteristics. Combining classification and hedonic regressions is not a novelty in itself, but in what follows the focus in on explicit evaluation and aggregation of the impact that different regressors have on the overall quality adjustment. The purpose is to obtain an index which fits as closely as possible into the traditional index number construction practices and at the same time makes full use of the hedonic quality adjustment procedure.

3.1 Classification

Location, type of building and number of rooms are the most fundamental characteristics of the dwelling, since they cannot be changed afterwards at all or only at large cost. It is also with respect of these characteristics that prices vary most. The regional stratification used was determined so as to form interpretable geographical entities with relatively similar price-level. The largest municipalities, for which separate index series are to be published were divided into two to four subregions by examining average prices of dwellings in 1995. On the other hand, smaller municipalities with few transactions were grouped together on regional basis. Within each region the dwellings were divided by type and number of rooms as follows

apartments in blocks of flats			apartments in terraced houses	
1 room	2 rooms	at least 3 rooms	1 or 2 rooms	at least 3 rooms

73,7 percent of the total price variation in the data for the year 2000 is between the cells suggesting that the adopted classification groups the observations into relatively homogenous groups.

After the data is classified, construction of a classical Laspeyres index straightforward. One proceeds by treating houses in each cell of this classification as perfect substitutes. Under this assumption, average price changes within each cell provide an unbiased estimate of price change and the index will be obtained by aggregating average prices across cells using Laspeyres' formula. We need only to agree on how cell price averages will be computed and how the cell quantities will be determined. In this case we use geometric price cell averages of square metre prices throughout. The rationale is that half-logarithmic regression modelling provides exact decompositions of within-cell geometric average price ratios into pure price and quality change components, and this feature is very attractive in what follows. Geometric averages are also recommended by HICP. The fixed Laspeyres quantities are the total floor area in the housing stock in each cell in 2000. Using weights based on stock rather than on the transacted dwellings in the base period is motivated by the fact that a house price index of old dwellings seeks to answer the question of how prices in the stock have developed on average. Purchasing a house is a very different concept than consuming housing services and thus weights based on transacted dwellings will have no clear meaning. With these conventions the house price index based on simply classifying the data is:

(3.1) $$Ind_0^t = \frac{\sum_{i=1}^{N} (\overline{floor\_area}_0^i * n_0^i) * \overline{p}_t^i}{\sum_{i=1}^{N} (\overline{floor\_area}_0^i * n_0^i) * \overline{p}_0^i}$$

where

N is the number of cells in the classification

$\overline{p}_0^i$ and $\overline{p}_t^i$ the average geometric prices at the base and the comparison period in cell i respectively

$\overline{floor\_area}_0^i$ and n_0^i the geometric average floor area and the number of dwellings in cell i in the base period

3.2 The Within-Cell Hedonic Quality Adjustment and its Decomposition

The above classification does not consider, among others, age of the house, floor area and micro-location, so price variation due to sample mix changes with respect to these characteristics will pass as price change in the index.

The quality adjustment strategy proposed is the following. A regression of standard type

(3.2) $$Ln(p_{itj}) = \boldsymbol{\beta}_t' \mathbf{x}_{itj} + \boldsymbol{\varepsilon}_{itj}$$

is specified either separately for each cell or for a larger section of the data including the necessary indicator variables to ensure that the sum of residuals for each cell are identically 0 by the properties of ordinary least squares (OLS) estimation. The sub-index i refers to the cell, t refers to the time period and j refers to the observation. p_{itj} is the price per square metre of floor area and $Ln(p_{itj})$ its natural logarithm. $\mathbf{x}_{itj}$ is the vector of characteristics, $\boldsymbol{\beta}_t$ is the vector of unknown parameters to be estimated and ε_{itj} is the statistical error term.

Denote the OLS-estimate of the parameter vector for cell i in the base period by $\hat{\mathbf{b}}_0^i$, and for the comparison period by $\hat{\mathbf{b}}_t^i$, the cell average vector of characteristics in the base period by $\overline{\mathbf{x}}_0^i$, the corresponding vector for the comparison period by $\overline{\mathbf{x}}_t^i$, and the geometric average prices in the two periods as $\overline{p}_0^i$ and $\overline{p}_t^i$ respectively. The following decomposition of the geometric average price ratio than is identically true:

$$(3.3) \qquad \frac{\overline{p}_t^i}{\overline{p}_0^i} = \exp\left[\hat{\mathbf{b}}_0^{i\,\prime}(\overline{\mathbf{x}}_t^i - \overline{\mathbf{x}}_0^i)\right]\exp\left[(\hat{\mathbf{b}}_t^i - \hat{\mathbf{b}}_0^i)'\overline{\mathbf{x}}_i\right]$$

The first term of the expression has clearly the interpretation of fraction of the price ratio due to quality difference of the sample mix at base period valuation of the characteristics. The second term can be interpreted is a price change due to changes in valuations, that is a "true price change". This type of decomposition introduced by Oaxaca (1973) is well known in the literature on wage discrimination, but to the author's knowledge is not commonly used in research related to hedonic indices. The quality adjusted price ratio is then

$$(3.4) \qquad \exp\left[(\hat{\mathbf{b}}_t^i - \hat{\mathbf{b}}_0^i)'\mathbf{x}_t^i\right] = \frac{\overline{p}_t^i}{\overline{p}_0^i \exp\left[\hat{\mathbf{b}}_0^{i\,\prime}(\overline{\mathbf{x}}_t^i - \overline{\mathbf{x}}_0^i)\right]} = \frac{\overline{p}_t^i}{\exp\left[\hat{\mathbf{b}}_0^{i\,\prime}\,\overline{\mathbf{x}}_t^i\right]},$$

since by the properties of OLS it is true that

$$(3.5) \qquad \overline{p}_0^i = \exp\left[\hat{\mathbf{b}}_0^{i\,\prime}\,\overline{\mathbf{x}}_0^i\right]$$

The quality adjusted cell average price in the comparison period, denoted as $\overline{p}_t^i(qa)$ is respectively

$$(3.6) \qquad \overline{p}_t^i(qa) = \frac{\overline{p}_t^i}{\exp\left[\hat{\mathbf{b}}_0^{i\,\prime}\,\overline{\mathbf{x}}_t^i\right]}\overline{p}_0^i$$

Denoting by N_t^i the number of observations in cell i in period t, the right hand side of (3.4) can be written also as

$$(3.4') \qquad \frac{\overline{p}_t^i}{\exp\left[\hat{\mathbf{b}}_0^{i\prime}\,\overline{\mathbf{x}}_t^i\right]} = \frac{\exp\left(\frac{1}{\mathrm{N}_t^i}\sum_{j=1}^{\mathrm{N}_t^i} Ln(p_{itj})\right)}{\exp\left(\frac{1}{\mathrm{N}_t^i}\sum_{j=1}^{\mathrm{N}_t^i}\hat{\mathbf{b}}_0^{i\prime}\,\mathbf{x}_{itj}\right)} = \left(\prod_{j=1}^{\mathrm{N}_t^i}\frac{p_{itj}}{\exp(\hat{\mathbf{b}}_0^{i\prime}\,\mathbf{x}_{itj})}\right)^{\frac{1}{\mathrm{N}_t^i}}$$

Since $\exp(\hat{\mathbf{b}}_0^{i\prime}\,\mathbf{x}_{itj})$ in the last term of (3.4') is a consistent price prediction of what a particular dwelling actually sold in period t would have been in the base period, it follows that $\frac{p_{itj}}{\exp(\hat{\mathbf{b}}_0^{i\prime}\,\mathbf{x}_{itj})}$ is a consistent prediction of the price ratio of this dwelling between the two periods and corresponds to a matched-pair price ratio with the difference that the denominator is not an actually observed price but a price prediction. Now it is clear, that the quality adjustment price ratio (3.4) has the interpretation of geometric average of matched-pairs price ratios and is completely analogous to what statisticians would normally compute at cell level in the classical index construction set-up.

It is worth noting that equation (3.2) is linear with respect of the coefficients $\boldsymbol{\beta}_t$ but not necessarily with respect to characteristic vector $\mathbf{x}_{itj}$, some of whose elements may be polynomials or other transformations of certain "basic" variables, e.g. age and squared root of age, which is our case as shall be seen later on in the text. In such situations the validity of results (3.3) – (3.6), especially the interpretation of (3.4) as an average of matched-pairs ratios, requires that the average $\overline{\mathbf{x}}_t^i$ is computed separately for each element, ignoring possible functional dependencies between the elements.

To clarify the point, let explicitly denote the characteristics vector $\mathbf{x}_{itj}$ as an appropriately defined function of some functionally unrelated characteristics vector $\mathbf{y}_{itj}$ that is $\mathbf{x}_{itj} = \mathrm{g}(\mathbf{y}_{itj})$. In the above discussion it is assumed that the average characteristics vector $\overline{\mathbf{x}}_t^i$ is computed over the observations in cell i in period t as $\overline{\mathbf{x}}_t^i = \frac{1}{N_t^i}\sum_{j'=1}^{N_t^i}\mathbf{x}_{itj} = \frac{1}{N_t^i}\sum_{j'=1}^{N_t^i}\mathrm{g}(\mathbf{y}_{itj})$. A natural notation for the last term is $\overline{\mathrm{g}(\mathbf{y}_{itj})}_t^i$. This is emphasised, because there is another possibility. Let $\overline{\mathbf{y}}_t^i$ denote the average of the "basic characteristics" $\mathbf{y}_{itj}$ in cell i for period t. $\overline{\mathbf{y}}_t^i$ is interpreted as the characteristics of the average property sold. It would be natural to ask what is the quality-adjusted price ratio at the "representative" point in the space of characteristics defined as $\mathrm{g}(\overline{\mathbf{y}}_t^i)$. If g(.) is non-linear, the answer to this question will differ from (3.4), because then $\overline{\mathrm{g}(\mathbf{y}_{itj})}_t^i$ will not in general equal $\mathrm{g}(\overline{\mathbf{y}}_t^i)$. However, according to Vartia and Koskimäki (2001), whose paper examines different quality adjustment alternatives, in many cases the difference will not have practical importance.

Another very important point is that decomposition (3.6) is not unique. The following equality is also identically true:

$$(3.7) \qquad \frac{\bar{p}_t^i}{\bar{p}_0^i} = \exp\left[\hat{\mathbf{b}}_t^{i\,\prime}(\bar{\mathbf{x}}_t^i - \bar{\mathbf{x}}_0^i)\right]\exp\left[(\hat{\mathbf{b}}_t^i - \hat{\mathbf{b}}_0^i)'\bar{\mathbf{x}}_0^i\right]$$

After some arithmetic manipulations it is seen that (3.7) implies the following quality adjusted cell average price for the comparison period:

$$(3.8) \qquad \bar{p}_t^i(qa)'' = \exp\left[\hat{\mathbf{b}}_t^{i\,\prime}(\bar{\mathbf{x}}_0^i)\right]$$

While the quality adjusted price representation (3.6) amounts to updating the base period prices by an evaluation of the price change of the observed comparison period characteristics mix, (3.8) is a direct evaluation of base period mix at comparison period prices. The term

$$(3.9) \qquad \frac{\bar{p}_t^i}{\exp\left[\mathbf{b}_0^{i\,\prime}\bar{\mathbf{x}}_t^i\right]}$$

certainly has a cell-level Paasche index interpretation and may seem inconsistent with the Laspeyres framework. However, another interpretation, expressed in the terminology of HCPI, is, that the variants at elementary aggregate level are constantly changing and then the quality adjustment procedure for the price of the comparison period variant mix is certainly not against the current principles of HCPI.

Using (3.6) rather than (3.8) is preferable in our case for the following reason. In the empirical part of the paper a quarterly index is computed, but a whole year is chosen as base period. Consequently there is much more data to estimate the base period coefficients than to estimate comparison period coefficients. The abundance of degrees of freedom allows using extensive set of location dummies covering separately areas where only few transactions per year occur. If one were to estimate regressions on quarterly basis, one would be restricted to much narrower set of location dummies to ensure that there are observations in each location for every period. With the chosen specification one is able to evaluate (3.9) and hence (3.6) for any subset of locations considered in the base period model that may occur in the comparison quarter's sample.

Another argument in favour of (3.6) over (3.8) is that the regressions for the comparison period do not have to be estimated. This is of great importance for official statistics where a production system should be as simple as possible. Still there is no restriction on the time structure of underlying parameters, using (3.6) simply does not require explicit estimates.

3.3 Evaluating the Impact of Characteristics on the Quality Adjustment

Evaluation of the effect on different characteristics on the quality adjustment and the index will greatly improve the transparency of the statistical procedure, facilitate its empirical evaluation and provide useful information for further analysis of the housing market.

An exact decomposition consistent with the discussion in section 3.2 is possible for index formulae having logarithmic representation, such as the Törnqvist formula. Here we discuss a more simple case for the log-Laspeyres formula defined in standard notation (see Vartia (1976)) as:

$$\text{(3.10)} \qquad \log\!-La_0^t \equiv \exp\Big(\sum_{i=1}^{N} w_0^i \ln\Big(\frac{\overline{p}_t^i}{p_0^i}\Big)\Big), \text{ where } w_0^i = \frac{p_0^i q_0^i}{\sum_{i=1}^{N} p_0^i q_0^i}$$

Now using (3.4) we have for the within cell quality adjusted index

$$\text{(3.11)} \qquad \log\!-La_0^t = \exp\Big(\sum_{i=1}^{N} w_0^i \ln\Big(\frac{\overline{p}_t^i}{\overline{p}_0^i \exp\Big[\hat{\mathbf{b}}_0^{i\,\prime}(\overline{\mathbf{x}}_t^i - \overline{\mathbf{x}}_0^i)\Big]}\Big)\Big)$$

$$= \exp\Big(\sum_{i=1}^{N} w_0^i \ln\Big(\frac{\overline{p}_t^i}{\overline{p}_0^i}\Big)\Big)\exp\Big(\sum_{i=1}^{N} w_0^i \hat{\mathbf{b}}_0^{i\,\prime}(\overline{\mathbf{x}}_0^i - \overline{\mathbf{x}}_t^i)\Big)$$

The first term of (3.11) is the pure classification log-Laspeyres index, and the second term is the explicit within-cell quality adjustment at aggregate level. One can group the estimated vector of characteristics valuation and the average characteristic vector into i.e. location (*L*) and size (*S*) components as follows:

$$\text{(3.12)} \qquad \hat{\mathbf{b}}_0^i \equiv \big[\hat{\mathbf{b}}_{0L}^i \hat{\mathbf{b}}_{0S}^i\big],\ \overline{\mathbf{x}}_0^i \equiv \big[\overline{\mathbf{x}}_{0L}^i \overline{\mathbf{x}}_{0L}^i\big],\ \overline{\mathbf{x}}_t^i \equiv \big[\overline{\mathbf{x}}_{tS}^i \overline{\mathbf{x}}_{tS}^i\big]$$

Then from (3.11) and using (3.12) one can decompose the quality-adjustment term into location and size components:

$$\text{(3.13)} \quad \exp\Big(\sum_{i=1}^{N} w_0^i \hat{\mathbf{b}}_0^{i\,\prime}(\overline{\mathbf{x}}_0^i - \overline{\mathbf{x}}_t^i)\Big) = \exp\Big(\sum_{i=1}^{N} w_0^i \hat{\mathbf{b}}_{0L}^{i\,\prime}(\overline{\mathbf{x}}_{0L}^i - \overline{\mathbf{x}}_{tL}^i)\Big)\exp\Big(\sum_{i=1}^{N} w_0^i \hat{\mathbf{b}}_{0S}^{i\,\prime}(\overline{\mathbf{x}}_{0S}^i - \overline{\mathbf{x}}_{tS}^i)\Big)$$

Naturally (3.13) can be extended to examine the aggregate effect on quality adjustment of each explanatory variable in the regression included.

Decomposition (3.13) does not hold exactly for the Laspeyres index, so either log-Laspeyres should be used[155] or (3.13) can be used as an approximation to evaluate the approximate effects of the quality adjustment factors on the Laspeyres index as follows. Define:

$$(3.14) \qquad \delta_0^t \equiv \frac{\ln(La_0^t)}{\ln((\log - La_0^t)}, \text{ so that } La_0^t \equiv \{\log - La_0^t\}^{\delta_0^t}$$

Then obviously, the quality adjustment of the Laspeyres index due to e.g. location differences can be evaluated by:

$$(3.15) \qquad \left\{\exp \sum_{i=1}^{N} \mathbf{b}_{0L}^{i}{}' (\overline{\mathbf{x}}_{0L}^{i} - \overline{\mathbf{x}}_{tL}^{i})\right\}^{\delta_0^t}$$

In our case the log-Laspeyres and the Laspeyres indices are so close, that (3.13) can be used directly without empirical problems.

[155] Log-Laspeyres index is always lower than or equal to the Laspeyres index with the same base year and weights (see Fisher (1922) and Vartia (1976)). In our case the difference between the two is negligible.

4.Empirical Results

4.1 Data Source

Research and planning work for renewing the House price index of Statistics Finland started in 2001 after Statistics Finland and the Finnish Tax Administration agreed, that information on transaction prices of apartments in blocks of flats and terraced houses will be provided on quarterly basis. The tax authority collects the prices in connection with the asset transfer tax, which is paid by the buyer and amounts to 1.6 percent of the price. The tax authority data has been available for statistical purposes since 1987, but only on an yearly basis and with delay of about 7 months, thus it could not be used for compiling a quarterly index. The index based on the method described in the following sections is already in use in Statistics Finland. The results reported here cover the period from 1987 till the end of 2000. This is the whole history of the tax authority data. The data of the taxation authorities covers ultimately all transactions of apartments, but on a quarterly basis it covers currently 2/3 share . This is because the buyers have two months time after the purchasing date to pay the tax and inform the authorities and there is some delay caused by information processing. In practice all transactions brokered by a real estate agent (those account for about two thirds of all transactions), will be reported with very little lag, since the real estate agent is responsible for the due payment of the tax and the standard practice is that the tax is paid and the tax form is filled at the moment the transaction takes place.

Table 4.1 below presents the estimated coverage of the data for the whole country and the major towns during the 3rd quarter of 2001. The estimates are based on the assumption that the total number of transaction is the same as during the 3rd quarter of 2000. The precision is good enough for the purpose of overall evaluation of the applicability of the data for constructing a quarterly index. The average coverage of the transactions in the new data seems sufficient with few exceptions. It should be kept in mind that the process is going on and improvement will be for sure achieved. The basic problem seems to be the clearly lower than the average coverage of the transactions in the last month of the quarter (September in the above case). This feature of the data will surely persist also in he future, although not so clearly as now. This suggests that weighing of the monthly observations might be needed in order to capture correctly the price development during the whole quarter. Because there was no "real-time" quarterly data from the tax authority, a pilot research was conducted to analyse the behaviour of the indices under different assumptions of the coverage of the actual quarterly data. This pilot work shall be shortly overviewed later in the text. Since there is no particular value of this research in terms of the central issues of the paper, the results presented further are based on the total data available since 1987 till 2000.

Table 4.1 The estimated coverage of the data during the 3rd quarter of 2001

Region	2001, 3:rd quarter		September 2001	
	reported transactions	%-share of total in 3:rd quarter of 2000	reported transactions	%-share of total in September of 2000
Whole Country	**11819**	**68,9**	**2093**	**35,9**
Helsinki	1807	70,1	233	24,6
Vantaa	568	75,3	118	44,9
Espoo	217	26,0	0	0,0
Porvoo	113	75,8	16	33,3
Tampere	959	87,3	174	49,9
Turku	697	68,4	135	39,2
Oulu	506	91,5	50	27,0
Lahti	380	82,1	72	44,7
Jyväskylä	323	77,8	35	30,4
Kuopio	232	58,3	10	7,0
Pori	136	43,9	16	17,0
Kotka	184	74,5	40	49,4
Kouvola	170	111,1	49	98,0
Hämeenlinna	223	95,7	70	88,6
Vaasa	200	81,3	32	35,6
Joensuu	246	79,6	58	49,6
Lappeenranta	200	92,6	46	73,0
Seinäjoki	29	16,1	1	1,9
Mikkeli	142	100,0	47	109,3
Rovaniemi	124	63,6	18	24,3
Rauma	131	76,2	32	60,4
Kajaani	111	76,6	15	44,1

Other important feature of the data is the extensive use of different registers. The information directly provided by the tax authority contains information only on the transaction price, the dwelling floor area and the municipality of transaction. Using the official apartment identification code prices are linked to other information such as type of building, number of rooms, year of construction and location (postal code and co-ordinates). The sources for this information are the taxation register of real estates, maintained by the tax authorities, and from the building and dwelling register maintained by the Population Register Centre. There are some problems with the use of all these registers concerning primarily new dwellings, for which information may be available only with delay, however long experience at Statistics Finland indicates, that what comes to old apartments the overall quality of the data is high.

In many cases the building in which the sold dwelling was situated is identified, but there is no certainty as to which one of several equally sized apartments was actually transacted. This feature of the data makes utilisation of repeat-sales information rather problematic. Although the results of

Meese and Wallace (1997) discussed in section 1.3 quite clearly show that the assumption of time-constancy of the parameters in repeat-sales and hybrid models are very likely to be violated for data set gathered over long time period, it would have been an interesting exercise to examine what results these methods would have provided.

4.2 Regression Estimation

In line with the discussion in chapter 3 the following regression equations for each region were estimated for the whole year 2000 data

$$(4.1) \qquad Ln(p_{ij}) = \beta_0 + \sum_{l=1}^{L_i} \beta_l A_{lij} + \gamma_1 (\text{floor\_area}_{ij}) + \gamma_2 \sqrt{(\text{floor\_area}_{ij})} +$$

$$\delta_1 age_{ij} + \delta_2 \sqrt{age_{ij}} + \sum_{k=1}^{3} \upsilon_k rooms_{kij} + \eta_1 TH_{ij} + \eta_2 (TH)_{ij} * (rooms3)_{ij} + \varepsilon_{ij}$$

Regressions are estimated separately for each region in the classification rather than for each cell (region X type of building X number- of- rooms class), because degrees of freedom for many cells are not enough to obtain reasonably stable estimates for the unknown coefficients. This means, that within each location the explanatory variables are restricted to have the same coefficients. For this reason the subscript *i* refers now to location rather than classification cell, *j* refers to the observation and the subscript for time period *t* is omitted, since in the estimation only the year 2000 data is used.

The general form of the regression model is of standard semi-log type. The dependent variable $Ln(p_{ij})$ is the price per square metre of floor area. The variables A_l , $l = 1.. L_i$ are postal code area indicators for the municipalities, which are separately examined, and municipality indicator variables for the rest of the regions. The variables $rooms_k$, k=1, 2, 3 are room-class indicators. Square roots of age and floor-area variables are included to capture non-linearity of the age and floor area profiles.

TH is an indicator for terraced houses and *TH***rooms*3 is an interaction, which takes value 1 if a terraced house apartment has 3 or more rooms.

The purpose of the model is to provide information for quality adjustment with respect to age, dwelling-floor area and micro-location of the dwelling. The room-number indicators as well as TH and *TH***rooms3* - indicators are included in the regression for technical reasons, to ensure by the properties of the OLS-estimator that the sum of residuals will be zero for all cells in the classification, since results (3.3)-(3.6) and the decomposition (15) hold exactly only if this is true. Obviously, they are strongly correlated with the floor-area, but given the large data, the estimated coefficients for the later are reasonable (see table 4.2).

Some further comments on the choice of the explanatory variables are also in place. First of all, a very short list of apartments' physical characteristics is used. This is because our experience is that type of building and construction year are very strongly correlated with the availability or absence of other characteristics. The age of the dwellings is used as an explanatory variable and since all observations are from the year 2000, the age coefficients can equally well be interpreted as construction-year coefficients and therefore they capture the effects of such "omitted" characteristics. There is inherent ambiguity in the interpretation of the age coefficients in regressions like that and different interpretations affect the quality adjustment differently. The problem is discussed in detail in the following section.

The register data contains information about basic amenities, but in Finland virtually all dwellings have amenities such as some form of heating, hot water, WC and shower. Sauna is a Finnish peculiarity, which is standard equipment in apartments built since the beginning of the 90's and not available in apartments of blocks of flats built in the 70's or earlier. Garage is present in terraced house apartments and is almost always bought separately in the case of apartment in block of flats. An important variable in the data, provided earlier by major real estate agents was the agents' evaluation of the overall condition of the flat, but such information can't be obtained from the register.

The use of extended set of location indicators is justified on the basis that relative differences in price levels by location reflect differences in characteristics difficult to incorporate in the model directly, such as availability of different services, transport connections, recreational activities as well as intangible factors such as the image of the area, which is usually result of a long and complicated socio-economic process. But there is also a potential problem with the location dummies. A change in the relative price level in an area, captured by the coefficient of a location dummy may reflect changes in the quantities of the unobserved characteristics rather than changes in the relative scarcity of the characteristic or changes in buyers' tastes. In case this should happen, it should be in principle viewed as a quality change. A good example is the level of noise. Suppose an airport is just built near a particular location. The level of noise is probably a negative factor, which tends to reduce house prices. One can argue that a quality adjusted index should not view this as a price fall, because physically the same apartment now provides worse housing services than before because of increased noise level. In practice such issues are very difficult to treat in a completely correct way.

Most likely there are omitted variables in our model, but they do not necessarily bias the index, since what we need is not unbiased regression coefficients, but only an unbiased prediction of the price of the average characteristics vector at base period prices (see equation (3.6))[156]. Of course, parameters of wrong size and sign will render the decomposition (3.13) non-interpretable in economic terms.

Using the total data for the year 2000 (58,566 observations) altogether 64 regressions were estimated with a total of 557 location indicators. The Table 4.2 summarises the results in a concise form giving to expositional simplicity more importance than to mathematical purity. The main points remain valid also in the detailed report provided in the Appendix.

Table 4.2 Average estimated coefficients

Variable	Average coefficient*	Average t-value*
floor area	0.0067	3.49
square root of floor area	-0.1595	-4.70
Age	0.0111	4.29
square root of age	-0.1797	-6.44
1 room indicator	0.0328	1,42
at least 3 rooms indicator	0.0219	1,16
terraced house indicator	0,1106	4.18
Terraced house with at least three rooms interaction	0,0227	0,93
*Average values weighted by number of observations in each separate regression		

The average R^2 statistic is 38 %, meaning that the models explain on average 38 percent of the within-regional price variation. The statistic is not high, but two things must be noted. First, the regional classification alone already captures 70 percent of the total price variation in the data, so the classification by region and the regional regression together capture over 80 % of the total price variation. Second, the location indicators (not reported in the table because of their very large number) as well as the floor-area and age variables taken together are statistically highly significant, as expected.

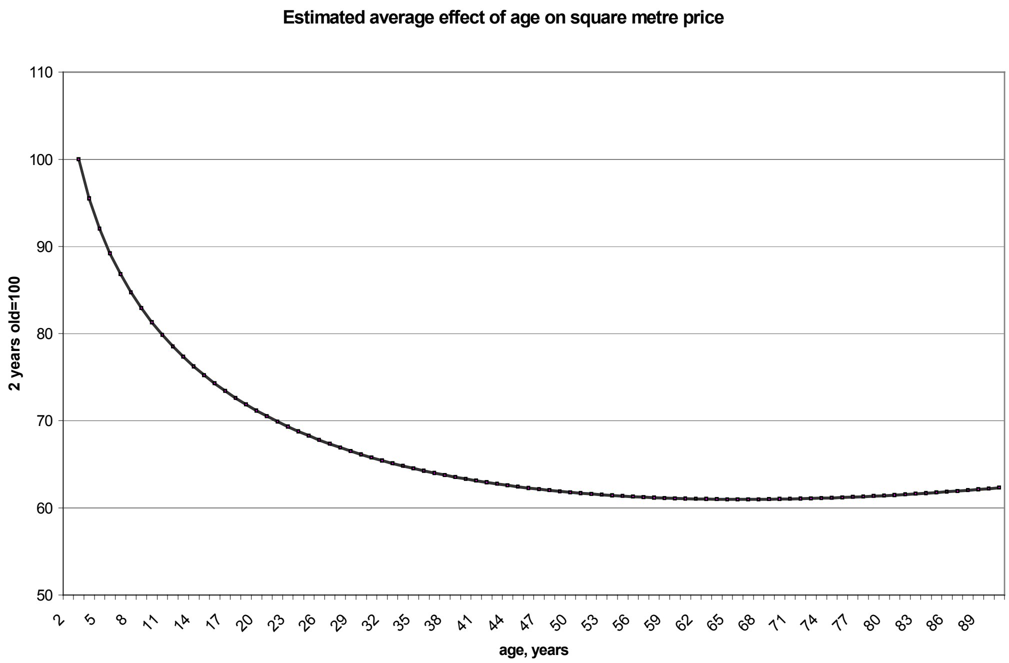
[156] This point is worth a short comment, because it seems to be sometimes misunderstood. Suppose the true equation is $y_i=\beta x_i+\gamma z_i+\varepsilon_i$ and there is a relation between z and x of the form $z_i=\delta x_i + \omega_i$ (i.e. the variables x and z are correlated) and z is omitted from the regression. The regression equation to be estimated is then $y_i=(\beta+\delta)x_i+(\varepsilon_i +\omega_i)$. Under standard assumptions the conditional expectation of $E(y_i|x_i)=0$ and OLS estimator will be unbiased for the "reduced form" parameter $(\beta+\delta)$. Consequently an unbiased prediction of y_i given x_i will be obtained. It is a different matter, that the estimated coefficient of x has no structural interpretation and is in this sense biased.

Although the coefficients of the room-number indicators are insignificant, they are kept in the model to ensure that the sum of residuals for each room- and type-of-building class is 0, so that the above presented price/quantity decompositions hold as identities.

The above diagram presents the effect of size on prices. The curve is evaluated using all size-related variables in the model, both the room-number indicators and the estimated dwelling floor area polynomial. It appears that square metre prices fall rather quickly with dwelling floor area. The square metre price of a 60 m^2 flat is 75 % of the square metre price of a 20 m^2 flat. This feature can be understood both from the technological and demand side. It is clear, that construction of large flats has economies of scale over small ones. On the demand side, Finland's single-person households account for 37 % of all households and the share is still growing, so demand for small apartments is high. The trend becomes flat for large, over 100 m^2 dwellings.

4.2 Interpretations of the Age Profile: Depreciation versus Vintage Effect

The diagram below presents the estimated average effects of age on prices of dwellings. According to our estimates the price initially falls with age at a yearly rate of 4,5 percent but the rate decreases fast and a 10-year old dwelling will lose only 1,8 percent of its value within a year. The age effect on prices is practically zero for 40 year old dwellings and even turns slightly upwards for over 70 years old dwellings. The fast decline of the price for rather new dwellings may partly reflect extrapolation, since there are few observations of only few years old dwellings in the data. This will not cause serious problems, since the average age of sold dwellings is typically much higher.



The interpretation of the age profiles is of great importance. There are two distinct simultaneously acting reasons explaining the shape of the age profile. The first is that if repairs do not (on average) offset depreciation, then, ceteris paribus, an older home will be in worse condition than a newer one. Another factor affecting the shape of the profile is that newer homes presumably embody better planning and construction technology and probably are on average better equipped. This way of thinking gives also a reason why for old enough houses the age profile eventually turns up[157] - they represent certain historically valuable architectural style and thus have "museum value". In the text the value of a house associated with the architectural style and construction technology of a particular time period is referred to as vintage effect as opposed to depreciation effect associated with physical ageing.

[157] Another argument is that for old houses the value increase of the land exceeds the decline of value due to depreciation. This is not a valid point here, since the location indicators in the model control for land value differences

For the purposes of quality adjustment of the index separation of the vintage and the depreciation effects would be desirable, since they have different implications for the quality adjustment. The depreciation effect implies, that if the average age of the dwellings in the sample increases, then the index should be adjusted upwards to reflect the decline of the average quality mix in the data. On the other hand, an increase in the average construction year of the sample would usually[158] call for downward adjustment of the index, because the sample mix embodies newer and better technology.

The problem is that separation of the depreciation and vintage effects is not possible in cross-section data, because the construction year, the selling period and the age are altogether by definition in perfect linear relation. An explicit recognition of the problem and an attempt to solve it by combining cross-section and panel data is provided by Englund, Quigley and Redfearn (1998). Their argument goes as follows. The starting point is specification in line with the hybrid model of Quigley (1995) discussed earlier. The log-price equation is specified in equation (6) in the quoted paper, presented here in notation consistent with the one used in earlier chapters as

$$(4.2) \quad \mathrm{p}_{it} = \mathrm{P}_t + \boldsymbol{\beta}\mathbf{x}_{it} + \beta_y \mathrm{YR}_\mathrm{i} + \beta_\mathrm{d}\mathrm{AGE}_{it} + \zeta_i + \varepsilon_{it}$$

$\mathrm{YR_i}$ is the construction year of dwelling i in the data, AGE_{it} is the age of the dwelling at time period t, and the vector $\mathbf{x}_{it}$ contains the other explanatory variables, $\mathrm{P_t}$ is the time-dummy for the sales period, ζ_i is the dwelling-specific term and ε_it the error term. The authors suggest, that the perfect linear dependence between construction year, the selling period and age is solved by the following. They define a new "error" term as

$$(4.3) \quad \gamma_{it} \equiv \beta_\mathrm{d}\mathrm{AGE}_{it} + \zeta_i + \varepsilon_{it}$$

and then estimate from the repeat-sales sub-sample

$$(4.4) \quad \mathrm{p}_{it} = \mathrm{P}_t + \boldsymbol{\beta}\mathbf{x}_{it} + \beta_y \mathrm{YR}_\mathrm{i} + \gamma_{it}$$

Then using residuals of the regression (4.4) as estimates of γ_it and the definition (4.3) they continue to estimate the AGE_{it} coefficient (that is the depreciation rate) from:

$$(4.5) \quad \gamma_\mathrm{it} - \gamma_{\mathrm{i}\tau} = (\beta_\mathrm{d}\mathrm{AGE}_{it} + \zeta_i + \varepsilon_{it}) - (\beta_\mathrm{d}\mathrm{AGE}_{i\tau} + \zeta_i + \varepsilon_{i\tau}) =$$
$$\beta_\mathrm{d}(\mathrm{AGE}_{it} - \mathrm{AGE}_{i\tau}) + (\varepsilon_{it} - \varepsilon_{i\tau}) = \beta_\mathrm{d}(\mathrm{t} - \tau) + (\varepsilon_{it} - \varepsilon_{i\tau})$$

[158] If the average construction year is very old, then the "museum value" may reverse the argument.

Unfortunately, there is a flaw in the procedure. The assumption that omitting the age variable from the regression will leave the depreciation effect in the error term is incorrect, because age is not uncorrelated with the included variables but on the opposite, it can be expressed as a linear combination of selling time and construction years. Omitting the AGE_{it} from the regression (4.2) will not leave the depreciation effect in the modified error term γ_{it}. The depreciation effect will be completely augmented in the estimated coefficients of the time dummies and the construction year.

The following discussion clarifies the point, which should be quite obvious. In their paper Englund et al. use data from 1981 first quarter to 1993 and construct a quarterly index using time dummies to denote the selling quarter of a house. The dummies are formed in a more complicated way than normally to identify precisely the sale date within a quarter (footnote 2 in the quoted paper). For clarity of exposition, let's suppose here that the sale quarter dummies are formed in the usual way, that is a quarter dummy takes value 1 if the sale is during the quarter and 0 otherwise. Suppose also that the age is measured in years and quarters. It is identically true that AGE_{it}=yyyy/q-YR_i, where yyyy/q denotes the year and quarter of sale. Substituting this expression into (4.2) and rearranging it one obtains

$$\text{(4.2')} \qquad p_{it} = P_t + \beta_d * \text{yyyy/q} + \boldsymbol{\beta}\mathbf{x}_{it} + (\beta_y - \beta_d)YR_i + \zeta_i + \varepsilon_{it}$$

For all dwellings sold in a particular quarter the term β_d*yyyy/q is constant and can be denoted as A_t and the parameter $(\beta_y-\beta_d)$ as β_v. Now what equation (4.4) actually estimates is (4.2'), which can be written also as

$$\text{(4.2'')} \qquad p_{it} = (P_t + A_t) + \boldsymbol{\beta}\mathbf{x}_{it} + \beta_v YR_i + \zeta_i + \varepsilon_{it}$$

The estimated coefficients of the time dummies are contaminated by the depreciation effect, which may bias the index if depreciation is significant. The estimated coefficient of construction year might appear to represent the "pure" vintage effect, but this is incorrect. Pure vintage effect could be estimated only if it were possible to evaluate the depreciation, which is not the case. Regressing the differences of the residuals of (4.2') for dwellings sold more than ones on the time between sales as in (4.5) does not estimate the depreciation, but can be rather viewed as a test for time-constancy of the parameters, which is assumed throughout in this setting. Englund et al. estimate the annual "depreciation" rate to be between 0,0023 and 0,00993, the estimates being statistically highly significant in all but one case. The result can be interpreted as evidence against the assumption of time-constant regression parameters.

The only plausible way to separate vintage and depreciation effect remains use of outside information such as expert judgement. Since such information is not easily obtained and its quality will be difficult to evaluate, one is left with the question of which factor is primary in determining the age profile. Most papers include age as explanatory variable in a model specification of the type (4.2), which means deciding implicitly in favour of depreciation. On a-priori grounds the choice is not obvious at all. Assuming that the downward sloping age-profile reflects depreciation means that one views the ageing of the stock as quality deterioration which may or may not be the case, depending on the renovation and repair activity. In Finland at least, most privately funded dwellings are regularly repaired and improved and are on average in very good technical condition. On the other hand newer vintages presumably take advantage of better planning and building techniques and for sure have more and better equipment than older ones. As mentioned earlier in Finland for example saunas are available in almost all apartments of newer vintages. Interpreting the age-profile as reflecting primarily vintage rather than age effect in the Finnish case is probably the choice closer to reality and this is how it is treated here in the quality adjustment procedure. There is no reason to interpret the growth of the average age in the Finnish stock of dwellings as indicator of housing quality deterioration.

The interpretation of the estimated age profile as reflecting vintage effects is achieved through the following simple variable transformation (re-interpretation). When computing the index for a particular period age is not calculated as physical age at the time of sale but as (2000-construction year)[159].

To understand more clearly the practical importance of this discussion, lets go through a simple example. Suppose that the quality adjustment of the index due to "age" is to be computed between 1990 and 1995. Suppose further that the dwellings sold in 1990 were on average built in 1975 and the ones sold in 1995 were on average constructed in 1980. Using the results from table 4.2 the

[159] This does not cause any problems for computing the index for periods up to 2001. From 2001 onwards there is the possibility that some sold flats are build after 2000 and then the square root of age cannot be calculated for them. In such cases the variable (2000-construction year) is set to 0. Since such cases will be very rare in the following 2-3 years, their effect on the index is negligible. Statistics Finland will regularly update the index and the estimated regressions in 3-4 year time periods, so errors will not accumulate. It might appear, that replacing age and its square root with construction year and its square root and re-estimating the regressions would be a clearly better solution, but it is not necessarily so. Of course in the latter case there will be no problems to perform the mathematical operations for dwellings built after 2000, but the danger of extrapolation is present.

quality adjustment of the index change between 1990 and 1995 due to difference in average construction year will be approximately

$\exp\{(0.0111(2000-1975) - 0.1797\sqrt{2000-1975}) - (0.0111(2000-1980) - 0.1797\sqrt{2000-1980})\}$ that is about 3.9 percent downward correction reflecting that in 1995 the dwellings sold were of newer vintage and hence better and ceteris paribus more expensive. For comparison, suppose now that in both periods the average construction year were the same, 1975. Under the current interpretation of the age profile as reflecting vintage effects, the quality adjustment in the index will be zero. However, the physical age of the 1975-built dwellings in 1990 is 15 years and in 1995 20 years. If the age profile were interpreted as reflecting depreciation, then in this situation the quality adjustment would be computed as $\exp\{(0.0111(15) - 0.1797\sqrt{15}) - (0.0111(20) - 0.1797\sqrt{20})\}$, which would result in about 5.4 percent ***upward*** correction of the index, reflecting that the 1975-built dwellings got older in 5 years and presumably in worse condition.

5 The Indices

5.1 Some Clarifications

The indices for the period 1987-2000 were estimated using the total data and aplying both the classification and using a within-cell quality correction method. Aggregation of the index was done using both the Laspeyres formula (3.1) and the log-Laspeyres formula. Since there was practically no difference in the results, the official index of Statistics Finland uses the log-Laspeyres formula, for which the decompositions (3.11) and (3.13) hold exactly.

When deciding the classification care was taken not to have classes with systematically less than 5 observations, but on the other hand it was considered unnecessary restriction of the classification to require that each cell has at least five observations in every period. As a result in some (but few) cases there are less than 5 observations or even no observations at all in some cells. The effect of two alternative ways of treating cells with few observations was examined. First, the price change in a cell was imputed by using the estimated change in adjacent cells. For example, if in a particular region there happened to be less than 5 prices of 2-room apartments in blocks of flats, the quarterly change for the blocks of flats in the region is computed using only other than two-room apartments. The cell-level index for two-room apartments is updated by the estimated quarterly change. This amounts to assuming that the price development in the cell with few observations follows the same trend as the most likely “substitutes”. In some cases this assumption may not work very well since it is not clear how good substitute a single-room and two- or three-room apartments are. For this reason an alternative procedure was examined. The cell price indices were computed normally for all situations where there was at least one observation in the cell for the quarter. If there were no observations, it was assumed that the price in the cell remained unchanged from the previous quarter and the cell-level index was left unchanged. This may seem to bias the index, but in author's view it is not the case if the cell is not empty all the time. With a base index, as in this case, the whole price change between the base and the comparison period will be correctly estimated any time there are observations in the cell. What will go wrong are the estimated changes between periods with not enough observations. Since cells where few observations occur are also cells with small weight in the index, the difference in the results of the two procedures becomes negligible at rather low aggregate levels.

Another aspect that was studied was the difference of the coverage of the data for different months in the quarter. The data received immediately after the end of a quarter under-represents transactions concluded during the last month of the quarter. The appropriate treatment is to give month weights to the observations. The procedure may be problematic, if the coverage in the last month is very low, since then a single observation may have considerable effect on the calculations. To study the effect of weighting random samples from each quarter drawn. The sample was 60 % for the first and second month and 20 % for the last month of the quarter. Then the indices were estimated first by giving a weight of 3 for the last month of the quarter and 1 for the other two months and then without weighting. The difference in the procedure did not bring about any significant change in the overall picture as compared to a benchmark case, where a flat 60 % sample per quarter was drawn.

The results presented in what follows are based on all observations for the years 1987-2000, so the above discussion is of little relevance for them, but is important for the "real time" computations. To conform the history with the actual practice now already established, the calculations are done using the currently "official" procedure at Statistics Finland. Price changes for cells with less than 5 observations are imputed and observations from different months are weighted by month weights in the calculation of cell level averages of prices and quality characteristics. The month weights are derived from the average number of transactions in each month in the years 1995-2000 Regional level results are presented in the Appendix.

5.2 Behaviour of the Indices and the Hedonic Quality Adjustment

Diagram 5.1 The overall index for Finland 1987-2000

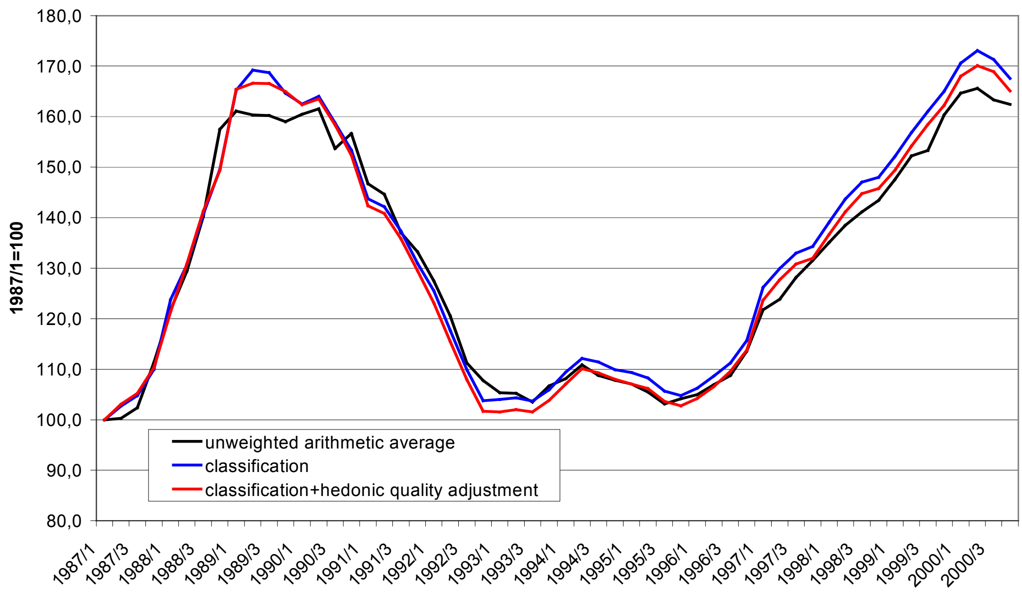
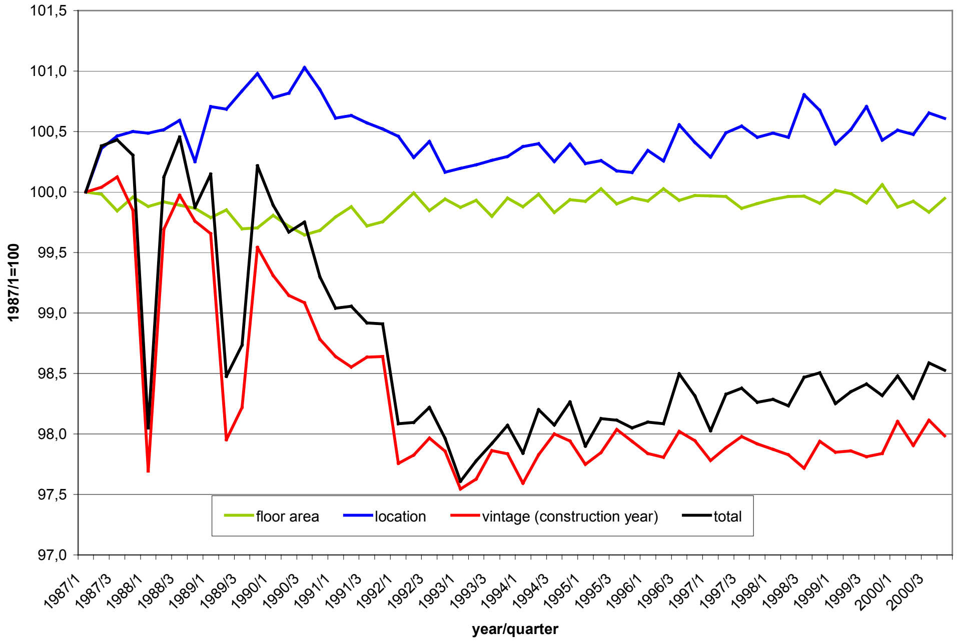


Diagram 5.1 presents results at the overall country level. The "classification" series is the index based on the classification described in section 3.1 and computed via log-Laspeyres formula (3.10) in the text. The unweighted average is computed by comparing the average square metre price in subsequent periods. The "classification + hedonic quality adjustment series" is the within-cell quality adjusted log-Laspeyres index, (3.11). At that level of aggregation even the simple average gives very similar picture of the long run price developments as the indices. Still, a closer look reveals a main problem of the simple average, which is not very serious in this particular case but is potentially dangerous. Both at the peak in 1989-1990 and during the price rise at the second half of the 90's the simple average series shows slower price movement than the indices. Obviously at that times the relative transaction volumes have shifted towards dwellings, where the square metre price is lower. Most likely this reflects shift in the demand towards relatively cheaper regions (locations).

The classification index takes account of the most important determinants of square meter price, region, type of building and number of rooms. This level of quality control seems to be quite enough at very aggregated levels. The within-class quality adjusted index does not change the overall picture. It is systematically below the classification index because by controlling for vintage effects (year of construction) it is taken into consideration that in the longer run newer and better equipped dwellings will enter the stock and improve the average quality of dwellings sold. A quality adjusted index should not interpret that feature as price rise. The point is clarified in diagram 5.2 presenting the overall quality adjustment and its components obtained via (3.13).

Diagram 5.2 The aggregate effect of quality adjustment 1987-2000, Finland, all dwellings



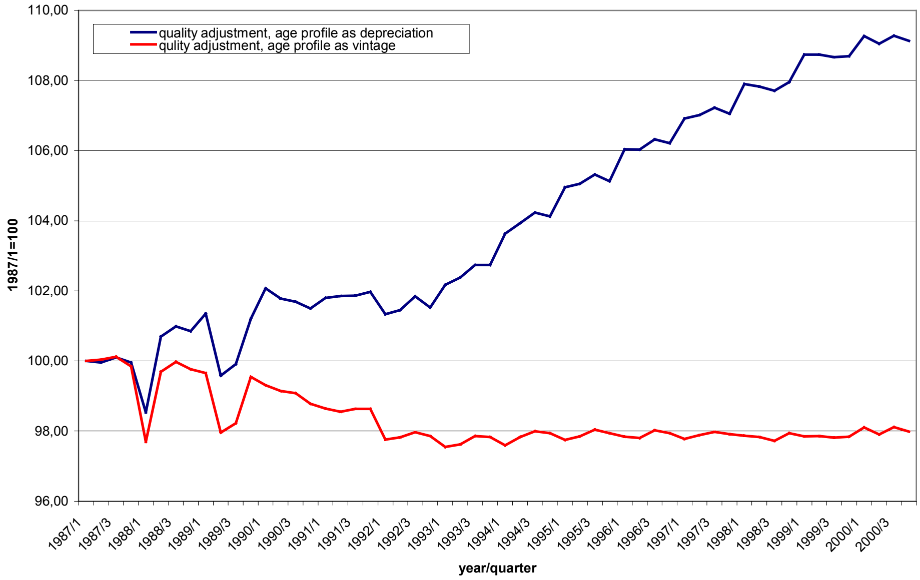
The most important factor affecting quality adjustment is the vintage effect. The high construction volumes at the end of the 80's are seen as a downward trend in the quality adjustment due to construction year: the average construction year became newer, which is interpreted in our method as indicating improvement in the average quality of the transacted dwellings. The trend practically ended after 1993. This observation is consistent with the very low construction volumes in the first half of the 90's and historically low construction volumes even after the economy recovered.

Although of very modest size, the quality adjustment due to micro-location of the dwellings supports economic intuition. As prices rose sharply at the end of the 80's the quality adjustment due to location goes upwards, reflecting shift of transaction volumes towards cheaper locations. The trend is reversed during the recession and turns up slightly again during the recent price upsurge.

The small negative quality correction due to floor area during the end of the 80's is also consistent with the logic of economic theory. The transaction volume has shifted towards smaller dwellings. The shift implies negative quality correction, since square metre prices are higher in smaller dwellings. The overall price of a smaller dwelling is, of course, lower than the price of a larger one, explaining why buyers would have to buy smaller flats when prices are very high.

The overall quality adjustment is negative most of the time and follows closely the vintage effect quality correction. Under the alternative interpretation of the age profile as reflecting depreciation rather than vintage effect discussed in section 4.2, the implications for the long term trend in quality adjustment would be the reverse, as illustrated in diagram 5.3.

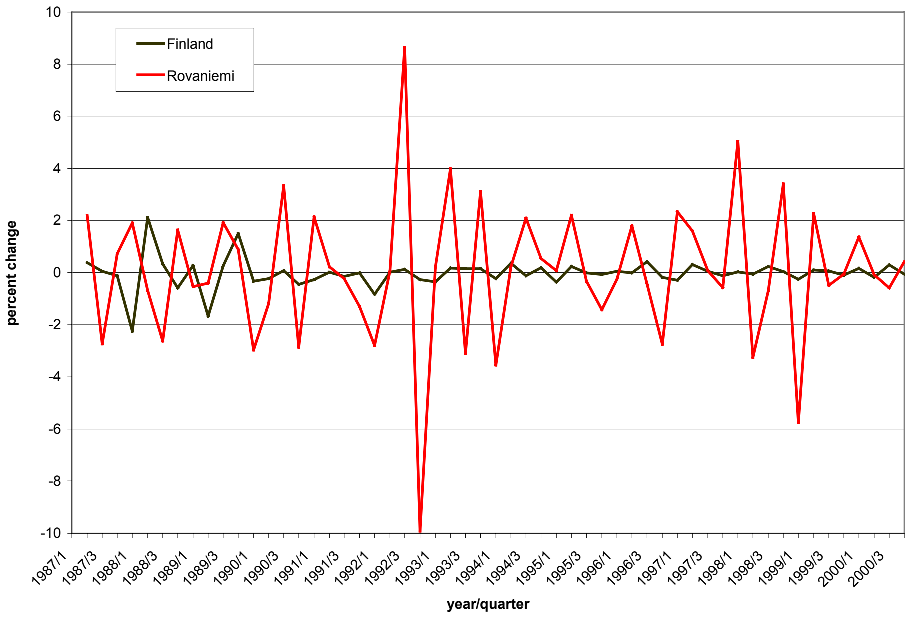
Diagram 5.3 Quality adjustment effects under alternative interpretations of the estimated age profiles



The red line in diagram 5.3 is the same as the quality adjustment due to vintage effects in diagram 5.2. The blue line is the quality adjustment generated under the interpretation of age profiles as reflecting depreciation. After the construction boom at the end of the 80's (seen with lag as a flat part of the line in the beginning of the 90's), new construction forms only a very small proportion of the dwelling stock. As a result each year the average age in the stock grows by almost 1 year. If the downward sloping age profiles reflect depreciation, the ageing of the stock means quality deterioration. In the Finnish case that interpretation leads to unduly strong upward quality correction of the index. It is a well known fact that dwellings in Finland are regularly repaired and even improved. The quality correction when age profile is interpreted as reflecting vintage effect is much smaller in absolute size, since at current construction volumes the average construction year (the average vintage) in the stock grows very slowly.

Although in the long run even the unadjusted average price changes provide a good picture of the price trend, the importance of the use of index formula and hedonic quality adjustment is better understood when viewing the house price index as a short term economic indicator. Then quarter to quarter changes rather than long term trends become important. In this respect unadjusted measures may be more misleading. The relative importance of quality adjustment is also much more important in small market areas than at overall country level.

Diagram 5.4 Quarterly changes in the overall hedonic quality adjustment in Finland and Rovaniemi



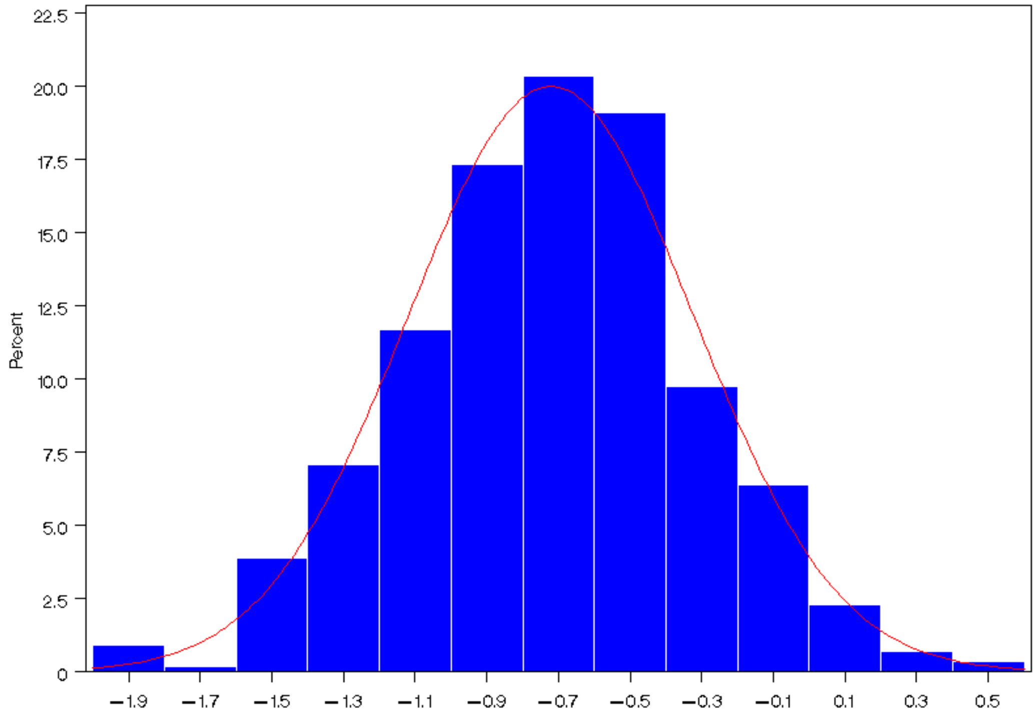
While at whole country level the quarter-to-quarter changes in overall hedonic quality adjustment is never larger in absolute size than 2 percent and is on average 0.3 percent, in Rovaniemi the average absolute quarter to quarter change is 1.96 percent and can be as large as 10 percent. Since 1 percent change in the quality correction between quarters implies approximately 1 percent change in the quarterly index, the strong influence of the hedonic quality adjustment for small markets such as Rovaniemi is obvious. As diagram 5.4 shows, a strong upward movement in the quality correction is usually followed by a strong downward movement and vice versa. This is due to the fact, that if in a particular quarter there is a random shift in the transaction volumes towards, say, high quality and high price dwellings, in the next quarter transactions of normal quality are likely to be observed. In the quarter-to-quarter quality adjustment change this is seen first as a peak downwards, to adjust for the better than the normal quality traded in the first quarter. This is followed by a peak upwards, since the return of transactions to dwellings of normal quality in the second quarter is, from the first quarter's prospective, a quality deterioration.

5.3 Precision of the Index

Precision in the index is evaluated by estimating the width of the 95 % confidence intervals for the quarterly percent change of the overall index and its sub-indices. The confidence interval widths are estimated by numerical simulation. The simulation set-up is the following. From the base year (2000) 565 samples with replacement are selected, each of the same size as the original data. Then from each sample the cell level average prices $\overline{p}_0^i$ and the regression coefficients for the cell level quality adjustment $\hat{\mathbf{b}}_0^i$ are computed. The vector of average prices and the matrix of regression coefficients from the k:th base year re-sampling are denoted as $\left[\overline{\mathbf{P}}_0, \hat{\mathbf{B}}\right]_k$. The quarterly change was arbitrarily chosen to be estimated between first and the second quarters of 1995. 565 samples with replacement of the same size as the original data set were selected from the data sets for the first two quarters of 1995 . Then cell level average prices $\overline{p}_{95/1}^i$ and $\overline{p}_{95/2}^i$ were computed. Denote the cell level average prices vectors for k:th re-sampling of the data in 1995 as $\left[\overline{\mathbf{P}}_{95/1}\right]_k$ and $\left[\overline{\mathbf{P}}_{95/2}\right]_k$ respectively. Now one has 565 independent "data sets" $\left[\overline{\mathbf{P}}_0, \overline{\mathbf{P}}_{95/1}, \overline{\mathbf{P}}_{95/2}\hat{\mathbf{B}}\right]_k$ From each data set index point number for the 1st and the 2nd quarter of 1995 and a quarterly percent change is computed at all levels of aggregation using exactly the same quality adjustment and aggregation procedures as explained in the text. In this way 565 independent values of the index point number are obtained for all sub-indices and the overall index. Since the average floor area and the number of dwellings in each class of the classification were evaluated independently using register information, these variables, which are used to evaluate the base period quantities, were treated as fixed (non-random) in the simulation exercise.

Diagram 5.5 presents the histogram for the distribution of the quarterly change between 1995 1st and 2nd quarters of the overall within-cell quality adjusted index

Diagram 5.5 The distribution of the quarterly change of the overall index between 1st and 2nd quarters of 1995



According to the simulation results the true quarterly change is between -1.46 and 0.03 percent with 95 % probability. The length of the 95 percent confidence interval is 1.49 percent. The original point estimate is -0.79, well within the confidence interval bounds and very close to the simulations' average, -0.72. The distribution is symmetric and very close to normal, the normal approximation to the confidence bounds is very good.

Regional results are provided in the Appendix. The distributions are generally symmetric and $\bar{x} \pm \frac{1}{2} h$, where h is the width of the simulated 95 percent confidence interval, provides a very good approximation to the simulation. The price distribution in the underlying population is so dispersed, that for example in a relatively large market areas such as Lahti, the width of the 95 % confidence interval for the quarterly change is over 9 percent. From the municipalities for which separate statistics are currently published the confidence interval is widest in Kokkola, 16 percent.

The confidence intervals were computed also for the quarterly changes computed from the classification index. They are in most of the cases only slightly wider than for the within-cell quality adjusted index. The author does not think that this fact bears any particular meaning for the desirability of the within-cell quality adjustment procedure, since its primary purpose is to remove potential bias in quarter-to-quarter changes and not necessarily to increase the measurement precision. Of course, the situation would be more complicated if it had turned out that the within cell quality adjusted index quarterly changes had much larger spread than the classification index. Then one possibly has to consider trade-offs between bias and variance, which is a complicated exercise.

It will be quite reasonable to assume, that the variance of the quarterly change depends only on the size of the data. If this is the case, then the estimates of the confidence interval lengths provided here are somewhat conservative, since the number of transactions per quarter is currently 20-25 percent higher than in 1995.

Conclusions

The methodology proposed in the paper is rather straightforward and transparent. The magnitude of quality adjustment due to difference in sample mix can be evaluated for each quality characteristic at each aggregation level. This makes comparison between a standard classification index and the within-cell quality adjusted index presented in the paper very clear. The trends in the quality adjustment components provide some interesting insights. The regression coefficients are allowed to change freely over time while their explicit estimation for each period is not necessary. This largely facilitates the use of the method in production of fast statistics, which are intended to be used as an economic indicator.

The paper does not provide empirical results based on modelling techniques using information on repeat-sales of the same dwelling. The reason is that precise identification of the repeat-sales in the data set used was problematic.

The econometric methodology used on the part of the quality adjustment decomposition is related to the literature on measurement of wage differences, which is a large and important field in econometric analysis of micro data, but the relation of its results to quality adjustment in index number construction does not seem to be generally known.

In abandoning the idea of time-invariant coefficients the paper bears some resemblance to the work of Meese and Wallace (1997). Another similarity is the finding that the long run price trends estimated using simple statistics are very similar to the trend exhibited by a sophisticated index. It seems that in the case of housing prices the role of quality adjustment is important primarily in estimating short-term price movements in relatively thin markets.

The only persistent trend of the quality adjustment found in the research is related to the age profile of housing prices. The shape of the age profile is a mix of vintage and depreciation effects. In view of their importance for the long-run index trend, it would have been highly desirable to separate the effects. Unfortunately, this does not seem to be possible on the basis of sample information, since time of sale, construction year and age are by definition in exact linear relationship and empirical models including all three are unidentified. The decision to interpret age profile as reflecting only vintage effects is simply a practical solution motivated by conditions in Finland.

The width of the evaluated confidence intervals of the quarterly index change suggests that in many cases quarterly changes at municipal and sub-municipal level should be interpreted with care. The precision of the overall index is good. A feature not examined in the paper is the precision of the estimates, when not all information on prices and transactions is available. Currently information about two thirds of all transactions concluded in a quarter are received by Statistics Finland immediately. Information for the rest is received later. The published index points and quarterly changes for the current year are revised and the final estimates are ready in the beginning of the following year.

The applicability of the developed method is evidenced by its use at Statistics Finland. It is of general nature and applicable to other cases where classification and following prices of representative items from each class is difficult. A good example would be construction of wage indices, where both the job characteristics and the qualifications of the worker affect wages and hence neither following the wages in same jobs nor wages of the same individuals is under all circumstances satisfactory. Hopefully the paper presents a good example of incorporating academic research with the practical needs of official statistics.

References

Articles

Bailey M. J., Muth, R. F. and Nourse H. O. (1963) A Regression Model for Real Estate Price Index Construction, *Journal of the American Statistical Association, Volume 58*, *933-942*

Calhoun Charles A., Chinloy Peter, Megbolugbe Isaac F. (1995) Temporal Aggregation and House Price Index Construction, *Journal of Housing Research, 6(3), 419-438*

Case K. E. and Shiller R. J.(1989) The Efficiency of the Market for Single Family Homes, *American Economic Review, 79(1), 125-137*

Diewert E. (2001) Hedonic Regressions: A Consumer Theory Approach, *conference paper, Joint ECE/ILO Meeting on Consumer Price Indices, Geneva, 2001*

Englund Peter, Quigley John M., Redfearn Christian L. (1998) Improved Price Indexes for Real Estate: Measuring the Course of Swedish Housing Prices, *Journal of Urban Economics, 44, 171-196*

Englund Peter, Quigley John M., Redfearn Christian L. (1999) The Choice of Methodology for Computing Housing Price Indexes: Comparisons of Temporal Aggregation and Sample Definition, *Journal of Real Estate Finance and Economics, 19:2, 91-112*

Fisher I.(1922) The Making of Index Numbers, Houghton Mifflin Company*, Boston*

Greene William H. (1997) Econometric Analysis (Third Edition), Prentice-Hall International, Inc., 1997

Griliches Z. (1971) Hedonic Price Indices for Automobiles: An Econometric Analysis of Quality Change*, Zvi Griliches (ed.) Price Indexes ad Quality Changes, 55-97*

Meese Richard A., Wallace Nancy E. (1997) The Construction of Residential Housing Price Indices: A Comparison of Repeat-Sales, Hedonic-Regression, and Hybrid Approaches, *Journal of Real Estate Finance and Economics, 14, 51-73*

Quigley John M. (1995) A Simple Hybrid Model for Estimating Real Estate Price Indexes, *Journal of Housing Economics, 4, 1-12*

Oaxaca R. (1973) Male-Female Wage Differences in Urban Labor Markets, *International Economic Review, 14, 673-709*

Rosen S. (1974) Hedonic Prices and Implicit Markets: Product Differentiation and Pure Competition, *Journal of Political Economy,82, 34-55*

Triplett J. (1983) Concepts of Quality in Input and Output Price Measures: A resolution to the User Value-Resource Cost Debate, In M.F. Foss, Ed., *The U. S. National Income and Product Accounts: Selected Topics, NBER Studies in Income and Wealth*, University of Chicago Press, 1983, pp. 269-311

Triplett J. (2001), Handbook on Quality Adjustment of Price Indices for Information and Communication Technology Products, *forthcoming, OECD, Paris*

Vartia Y.O. (1976) Fisher's Five-tined Fork and Other Quantum Theories of Index Numbers, *Theory and Applications of Economic Indices, Proceedings of an International Symposium Held at the University of Karlsruhe April-June 1976, 1978, 271-295*

Vartia Y.O., Koskimäki Timo (2001) Beyond matched pairs and Griliches-type hedonic methods for controlling quality changes in CPI sub-indices: mathematical considerations and empirical examples on the use of linear and non-linear hedonic models with time-dependent quality parameters, *Discussion Paper, Department of Economics, University of Helsinki*

Other Sources

Asuntojen hinnat 1970 – II neljännes 1985, *HI 1985:6, Tilastokeskus*

Asuntojen hinnat 2001 3. neljännes, *Asuminen 2001:8, Tilastokeskus*

Journal of Real Estate Finance and Economics, 14, Kluwer Academic Publishers, 1997

Appendix

Appendix 1.

Period	Average nominal quarterly interest rate. Commercial and savings banks.	10-year long-term interest rate on government bonds	**User cost (alt. 1)** Figures in EURO	**User cost index (alt. 2)** jul99=100	**User cost index (alt. 3)** jul99=100 **Assuming house price change = 2.0 %**	**Assuming house price change = 9.8 %**	**Rental equivalence** jul99=100
Jul.99		5.8	-1411	100.0	100,0	100.0	100.0
aug.99	7.8	5.8	-1411	100.0	100,0	100.0	100.0
sep.99	7.3	5.8	124	100.7	100,5	100.0	100.7
okt.99	7.3	5.8	124	100.7	100,5	102.5	100.7
nov.99	7.3	5.8	124	100.7	100,5	102.5	100.7
des.99	7.1	5.8	-313	106.1	101,0	102.5	101.0
jan.00	7.1	5.8	-313	106.1	101,0	105.0	102.6
feb.00	7.1	5.8	-313	106.1	101,0	105.0	102.7
mar.00	7.1	5.8	-1103	115.6	101,5	105.0	102.9
apr.00	7.1	5.8	-1103	115.6	101,5	107.5	103.5
mai.00	7.1	5.8	-1103	115.6	101,5	107.5	103.5
jun.00	7.4	5.8	-1336	127.2	102,0	107.5	103.8
jul.00	7.4	5.8	-1336	127.2	102,0	110.2	104.0
aug.00	7.4	5.8	-1336	127.2	102,0	110.2	105.3
sep.00	8.1	5.8	1958	129.4	102,5	110.2	106.0
okt.00	8.1	5.8	1958	129.4	102,5	112.9	105.6
nov.00	8.1	5.8	1958	129.4	102,5	112.9	106.4
des.00	8.5	5.8	953	132.0	103,0	112.9	106.3
jan.01	8.5	5.8	953	132.0	103,0	115.6	107.5
feb.01	8.5	5.8	953	132.0	103,0	115.6	107.8
mar.01	8.5	5.8	-259	136.9	103,6	115.6	107.8
apr.01	8.5	5.8	-259	136.9	103,6	118.5	108.2
mai.01	8.5	5.8	-259	136.9	103,6	118.5	108.4
jun.01	8.5	5.8	-562	140.0	104,1	118.5	108.6
jul.01	8.5	5.8	-562	140.0	104,1	121.4	108.6
aug.01	8.5	5.8	-562	140.0	104,1	121.4	108.9
sep.01	8.5	5.8	1024	138.8	104,6	121.4	109.7
okt.01	8.5	5.8	1024	138.8	104,6	124.3	109.9
nov.01	8.5	5.8	1024	138.8	104,6	124.3	110.6
des.01	8.4	5.8	1154	135.3	105,1	124.3	111.8
jan.02	8.4	5.8	1154	135.3	105,1	127.4	112.8
feb.02	8.4	5.8	1154	135.3	105,1	127.4	112.4
mar.02	8.0	5.8	-233	139.4	105,6	127.4	113.1
apr.02	8.0	5.8	-233	139.4	105,6	130.5	113.6
mai.02	8.0	5.8	-233	139.4	105,6	130.5	114.3
jun.02	8.0	5.8	-322	147.2	106,2	130.5	114.7
jul.02	8.0	5.8	-322	147.2	106,2	133.7	115.1
aug.02	8.0	5.8	-322	147.2	106,2	133.7	114.9
sep.02	8.5	5.8	2445	143.6	106,7	133.7	115.1
feb.02	8.5	5.8	2445	143.6	106,7	137.0	115.1
nov.02	8.5	5.8	2445	143.6	106,7	137.0	115.4
des.02	8.5	5.8	1198	131.1	107,2	137.0	115.6
jan.03	8.5	5.8	1198	131.1	107,2	140.3	117.1
feb.03	8.5	5.8	1198	131.1	107,2	140.3	117.4
mar.03	7.5	5.8	-9	121.2	107,8	140.3	117.8
apr.03	7.5	5.8	-9	121.2	107,8	143.8	118.8
mai.03	7.5	5.8	-9	121.2	107,8	143.8	119.0
jun.03	6.6	5.8	442	97.6	108,3	143.8	119.2
jul.03	6.6	5.8	442	97.6	108,3	147.3	119.3
aug.03	6.6	5.8	442	97.6	108,3	147.3	119.4
sep.03	4.8	5.8	748	91.9	108,8	147.3	119.6
okt.03	4.8	5.8	748	91.9	108,8	150.9	119.7
nov.03	4.8	5.8	748	91.9	108,8	150.9	119.9
des.03	4.4	5.8	200	91.9	108,8	150.9	120.1

Stefan Linz and Timm Behrmann[160]: Using Hedonic Pricing for the German House Price Index

1. Introduction

The objective of official price statistics is to measure what we call "pure" price changes, purged of the adulterating influence of changes in consumption patterns, types of goods and quality features. This essentially reflects the Laspeyres Principle of once defining a basket of goods and keeping it as constant as possible over a defined period of time.
Especially, the price of an item at two separate times can only be usefully compared if the *quality* of the item remains constant. If this is not the case, quality adjustment is undertaken in order to introduce the monetary value of an item's quality change into price observation. Hedonic methods constitute a specific quality adjustment technique. The hedonic method uses regression analysis to measure the influence of product features on the sale price. Thus price changes due to qualitative changes in certain features can be distinguished mathematically and purged from the pure price change which the price index is actually called upon to measure.[161]

In 2002 the Federal Statistical Office in Germany began an extensive programme for introducing hedonic techniques of quality adjustment. Figure 1 provides an overview of the stages in this schedule.
As a first step the hedonic method was introduced in June 2002 to the regular monitoring of prices for home computers.[162]
The second step for the German Federal Statistical Office was to evaluate the quality adjustment procedure hitherto applied to the consumer price index for motor vehicles. A hedonic price index was calculated parallel to the existing price index for new cars. Analysis demonstrated that, for new cars sold in Germany, quality changes due to technological progress are adequately indicated by the "traditional" method of quality adjustment which is so called "Option Pricing" in this case. No systematic deviations between the two indices were observed. The Federal Statistical Office has, therefore, not incorporated the hedonic method into its quality adjustment techniques for new cars, retaining instead the well-proven and significantly cheaper traditional approach.
The third step was to design a hedonic price index for used cars. In May 2003 a used car price index based on the hedonic approach was included in the consumer price index[163]. One year later, in May 2004, hedonic producer, import and export price indexes for selected data processing equipment have been implemented in the respective indices. Furthermore, work is currently progressing on hedonic price indexes for the categories "electrical home appliances" and "consumer electronics".

[160] Federal Statistical Office of Germany
[161] On hedonic price determination cf. inter alia: Griliches, Zvi (1971) Price Indexes and Quality Change, Studies in New Methods of Measurement, edited by Zvi Griliches for the Price Statistics Committee Federal Reserve Board, Harvard University Press, Cambridge, Massachusetts. Brachinger, H. W. (2002) Statistical Theory of Hedonic Price Indices, Working Paper, available in English at: http://www.unifr.ch/stat/en-home.php.
[162] Cf. Linz, Stefan and Eckert, Gudrun, Introducing hedonic methods in price statistics, available at: http://www.destatis.de/download/e/preise/hedonic-engl.pdf.
[163] Cf. Linz, Stefan and Dexheimer, Verena, Hedonic Methods of Price Measurement for Used Cars, available at http://www.destatis.de/download/e/preise/hed_used_cars.pdf

Figure 1: The Federal Statistical Office programme for implementing hedonic methods

Index position	*Project status*
Consumer prices for home computers	Hedonic index introduced in June 2002
Consumer prices for new cars	Evaluation completed in January 2003. Hedonic methods are not used to calculate this price index as no systematic deviations were observed between the techniques.
Consumer prices for used cars	Hedonic price index implemented in May 2003
Producer, import and export price indexes for electronic data processing equipment	Hedonic price indices implemented in May 2004
Consumer prices for owner occupied housing	Preliminary results in 2004
Consumer prices for electrical home appliances and consumer electronics	Work in progress

The present paper deals with the development of a hedonic price index for "Owner Occupied Housing" in Germany. It is based on the findings and experiences the Federal Statistical Office of Germany has gained within a pilot study initiated by the Statistical Office of the European Communities (Eurostat).

Primary aim of the 'Harmonized Index of Consumer Prices (HICP)' is to provide a measure of consumer price inflation on a comparable basis. Nevertheless, the coverage of the HICP is still different in most national Consumer Price Indices. Thus, for example, the HICP includes information on shelter prices that refer to actual rentals for housing paid by tenants, but actual prices for owner occupied residential property are still excluded up to now. The exclusion of Owner Occupied Housing is – mainly from the European Commission's as well as the European Central Bank's point of view – not satisfactory and leaves a gap in the coverage of household final consumption expenditure. Therefore a pilot study on the coverage of Owner Occupied Housing (development of 'House Price Indices') was launched by Eurostat in a limited number of countries (including Germany) in early 2002. The hedonic index computed for this project is the main focus of this paper. Overall, in spite of some difficulties, the main finding is that it seems feasible to compile a hedonic price index for owner occupied housing in Germany at least on a quarterly basis.

2. Data sources

In order to avoid the implementation of an extensive new survey on purchases of turnkey-ready houses, existing data has been used for this pilot study. For Germany, mainly the data of the construction price statistics as well as the building activity statistics come into consideration. Nevertheless, there is a lack of data in the field of "turnkey ready buildings" in Germany. In order to close this gap, appropriate data has been acquired from so-called advisory committees of land value ("Gutachterausschüsse für Grundstückswerte").

The main tasks of these advisory committees consist on the one hand in collecting data on the purchase of buildings and dwellings and on the other hand in estimating the value of houses and land. The advisory committees are mainly organised on a NUTS-3-level (district area level), partly on NUTS-4-level. All of them are independent committees and based on federal (not statistical) law. Furthermore, at present, they are not obliged to provide statistical data on the purchase of buildings. The main problem when using the data of the advisory committees was to coordinate and harmonise the different definitions of variables and concepts used by the committees for data collection. Nevertheless, it was possible to get a lot of detailed data which are at most part appropriate for an Owner Occupied Housing Index. A collaboration within several federal states has successfully been initiated and will be intensified in order to cover additional regions and federal states in future work.

The prices collected and reported by the advisory committees are real transaction prices (and not offer prices) at the time of the conclusion of the contract between the builder (building promoter)

and his client. For the pilot study, we could use data from five federal states out of 16 in Germany, which refer to the years 2000 to 2002.
The amount of data reported is quite different: The largest federal state supplies nearly 400 quotations per quarter of a year. The other states report much less data, sometimes only 20 quotations.
As characteristics are reported:

Physical characteristics
Type of dwelling/property type
Single-family house
Two-family/semidetached house
Terraced house
Free-standing house
Freehold flat in multi-storey buildings
Vintage
2000, 2001, 2002; if possible 1999 and 1998
Date of purchase (month, day)
Type of construction
Conventional ("Self builders"; not prefabricated)
Prefabricated
Furnishing and luxury elements (Sauna or swimming-pool included)
Cellar
Storing position for car (s)

Locational characteristics
State ("Bundesland")
County
City
Municipality
Part of municipality
Type of quarter
Downtown district
Outskirts
Rural regions
Location of building/dwelling in general
Simple/plain
Medium/average
Good/very good

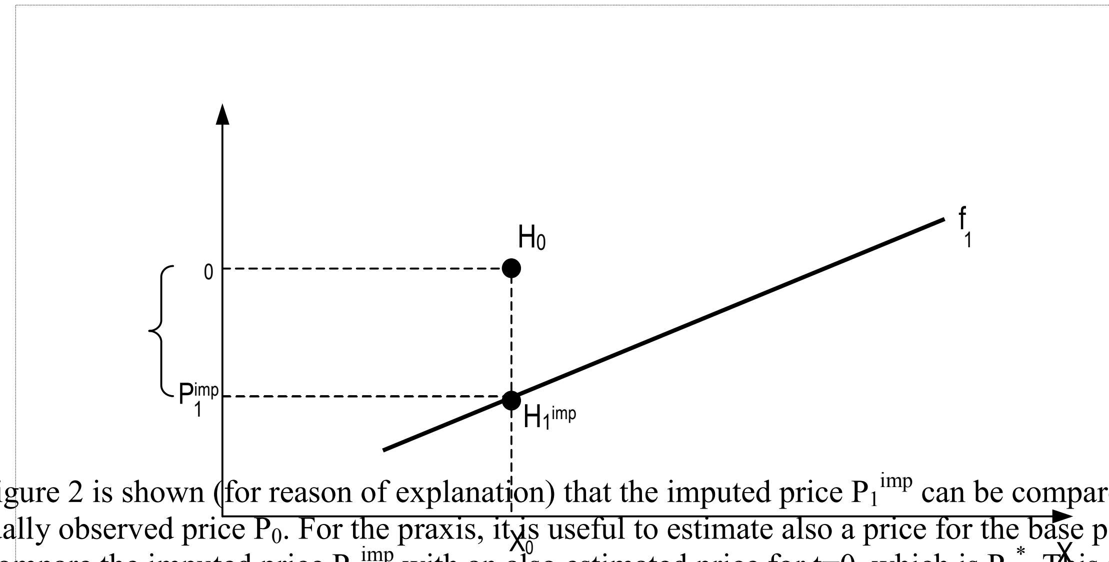
Generally price variables
Purchase price (real transaction price)
Size of the (developed) real estate/land/plot of land (in square metres)
Size of the living area (in square metres)
Proportionate price of the plot of land in relation to the total purchase price
If the proportionate price is not available: specification of the 'guiding private value' (in German: Bodenrichtwert"; in EUR/m²). The 'guiding private value' is approximately equivalent to the current market price of the location/plot of land.
Building/dwelling ready for occupancy at the point of time of purchase
Building/dwelling in stage/phase of building or planning

2. Imputation Method

Hedonic methods are well-suited to account for quality changes regarding the heterogeneous product 'house' respectively 'dwelling'. Each house or dwelling can be interpreted as unique, appearing only once on the market. In contrast to other products, a house observed in the base period, will never be exactly the same to be observed in a comparison period. Thus, the application of direct price comparison in order to calculate price indices is not possible.

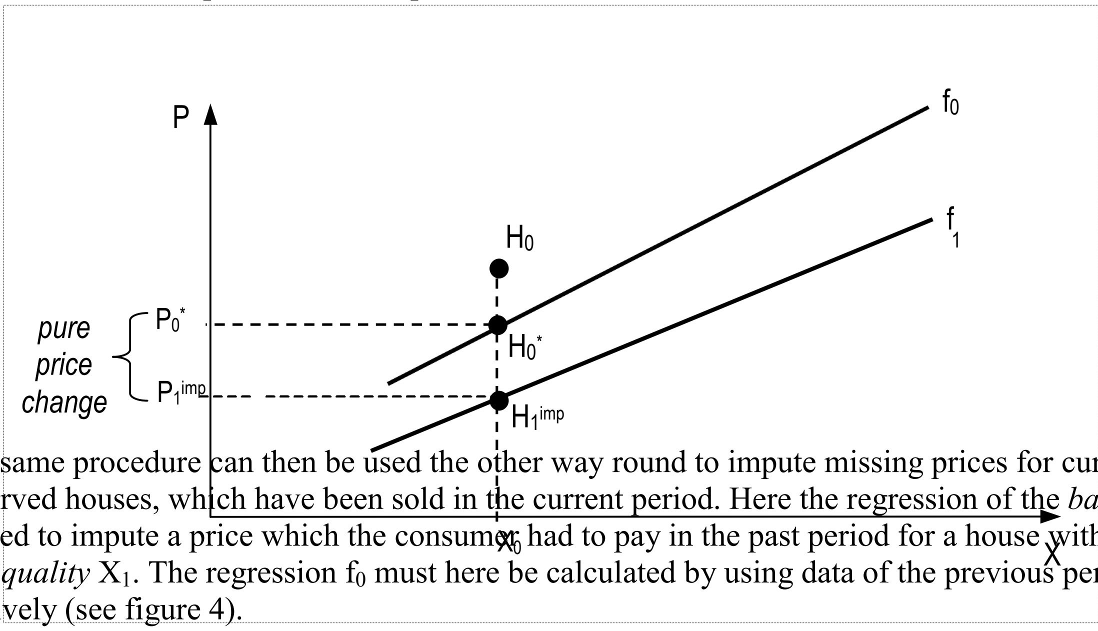
As hedonic method, the "Imputation Method" is used and missing prices are imputed by means of regression analysis. In figure 2 below, an example is shown for one House H_0. The house H_0 has been sold in the base period with a quality of X_0. The quality X is here meant as quality-vector, which include the values of *several* quality features. The principle is in the picture nevertheless shown for only one dimension of the quality.
We use the regression function f_1 to estimate the price, which the consumers have to pay under the current market conditions for a House with the old quality of X_0. It is important, that the regression f_1 is calculated by using data of the current period exclusively. Then the difference between P_0 and the estimated (or imputed) price $P^{imp}{}_1$ gives us the pure price change in the time period considered for the house H_0.

Figure 2: Imputing a current price for an old product



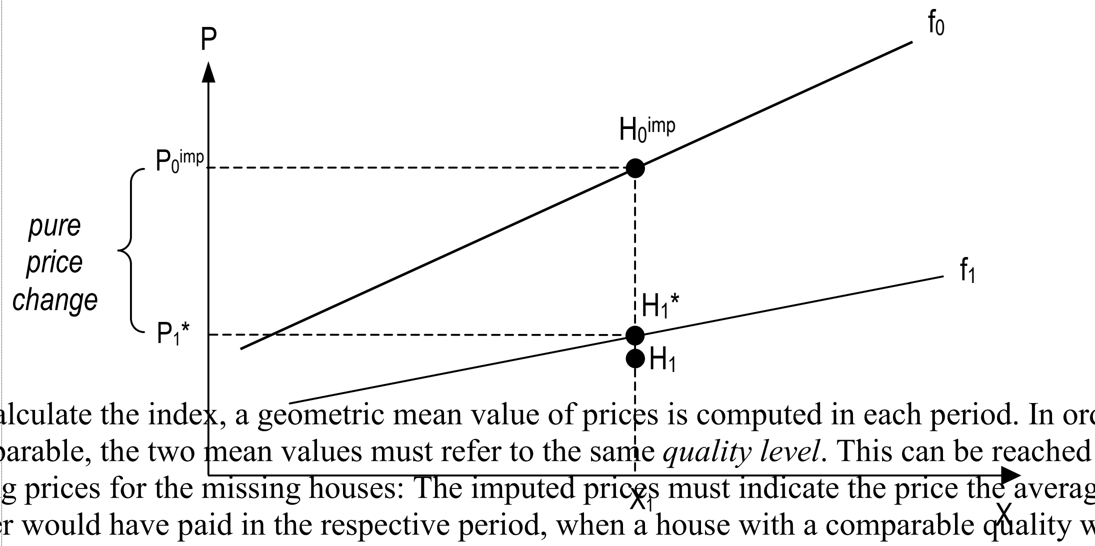
In figure 2 is shown (for reason of explanation) that the imputed price $P_1{}^{imp}$ can be compared to the actually observed price P_0. For the praxis, it is useful to estimate also a price for the base period and to compare the imputed price $P_1{}^{imp}$ with an also estimated price for t=0, which is $P_0{}^*$. This procedure is called *Double Imputation*.
The Double Imputation has the advantage, that outlier do not go into the index calculation. The directly observed price P_0 may be an outlier whereas the estimated price $P_0{}^*$ is the conditional expected value for this model. Thus, the robustness of the index is increased. (However, the regression functions are still calculated on the directly observed prices.)

Figure 3: Double Imputation for old products



The same procedure can then be used the other way round to impute missing prices for currently observed houses, which have been sold in the current period. Here the regression of the *base period* is used to impute a price which the consumer had to pay in the past period for a house with the *current quality* X_1. The regression f_0 must here be calculated by using data of the previous period exclusively (see figure 4).

Figure 4: Imputation method for new products



To calculate the index, a geometric mean value of prices is computed in each period. In order to be comparable, the two mean values must refer to the same *quality level*. This can be reached by imputing prices for the missing houses: The imputed prices must indicate the price the average consumer would have paid in the respective period, when a house with a comparable quality would have existed.

Figure 5: Whole sample with imputed and estimated prices

sub-sample	*base period*	*current period*
Houses from the current period	$P_0^{imp}(1)$	$P_1^*(1)$
	$P_0^{imp}(2)$	$P_1^*(2)$
Houses from the base period	$P_0^*(3)$	$P_1^{imp}(3)$
	$P_0^*(4)$	$P_1^{imp}(4)$
	geometric mean t=0	geometric mean t=1

$P_0^{imp}(n)$ *...imputed price to pay in time t for the quality of product n*

$P_t^*(n)$ *...estimated price to pay in time t for the quality of product n*

The geometric means of the two periods are then compared in order to derive the quality adjusted price development.

4. Regression Analysis

A first step regarding regression analysis is to address the issue of the ‘right’ choice of the functional form. Initially, the ‘linear model’ as well as the ‘log-linear form’ and the ‘log-log model’ were taken into consideration. After testing these various models, the log-log-form which is in its basic shape a multiplicative model and therefore taking into account interactions between variables, has proved to be most robust.

To calculate the index for single-family/two- family houses, the following regression function has been used:

(1) $$\ln(p) = \beta_0 + \beta_1 \cdot \ln(grund) + \beta_2 \cdot \ln(wohn) + \beta_3 \cdot \ln(BRW) + \beta_4 \cdot d\_haus + \beta_5 \cdot d\_keller + \beta_6 \cdot d\_stellplatz + \sum_{i=1}^{8} \beta_{6+i} \cdot d\_kreis\_i + \varepsilon$$

p … Price including plot of land in €
grund … size of plot of land [square metres]
wohn … size of living area [square metres]
BRW … class of ‘guiding private value’
d_haus … dummy for the type of dwelling: 1, when free-standing detached house

0, when semidetached/terraced house

d_keller … dummy for cellar/basement: 1, when cellar does exist
0, when cellar does not exist

d_stellplatz … dummy for storing position for car(s): 1=storing pos. for car does exist
0= does not exist

d_kreis_i … dummy for area type (1-9): 1, if area type = i, 0, if not

The figure 6 shows excerpts of outputs of the regression analysis for the first quarter of 2001 referring to detached/semidetached houses within the state "Lower Saxony".

Figure 6: Results of the regression analysis

variable	parameter	standard-error	t -value	variance inflation factor
intercept	9,681	0,31	31,09	0,00
d_haus	0,189	0,03	6,31	1,45
ln(grund)	0,118	0,03	3,62	1,89
ln(wohn)	0,316	0,07	4,63	1,47
d_kreis_1	0,748	0,10	7,27	1,20
d_kreis_6	-0,068	0,03	-2,15	1,35
d_kreis_9	-0,162	0,10	-1,65	1,09
d_BRW_0-50	-0,165	0,05	-3,40	1,27
d_BRW_100-199	0,143	0,02	6,19	1,41
d_BRW_200-299	0,291	0,06	4,62	1,32
d_keller	0,160	0,02	7,10	1,14
d_parkplatz	0,080	0,02	3,84	1,12

The site of the house plays an important role for the price. The site is represented by dummy variables for the area type as well as by the so called 'guiding private value' (Bodenrichtwert). The "guiding private value" is approximately equivalent to the current market price of the location/plot of land. It is classified, so that only the cross-sectional information goes into the regression analysis. An overall price increase of the guiding private value would not be measured as increase of quality. As a result, increasing prices due to higher land prices will be identified as "pure price increases" by the index. The house price index is therefore an index including the price of the house as well as the price of the plot of land.

5. Preliminary Results

The following results point out first experimental indices for turnkey ready buildings with regard to different federal states. In particular, for the East German federal states (Saxony, Saxony-Anhalt and Mecklenburg-Western Pomerania) the samples seem to be to small in part, because the results of the regression analysis for each of these federal states are unsatisfactory due to a lack of observations. Therefore, the data of all the Eastern federal states were pooled and aggregated in order to perform regression analyses which lead to suitable results. Thus, the following results show first price indices for Lower Saxony and East Germany (consisting of the three federal states mentioned above) only. The development of the index is quite similar for Lower Saxony and the Eastern Länder. There is a significant price increase in the 4th quarter of 2001 succeeded by a price decrease in the first quarter of the year 2002. This is probably due to the anticipation by the consumer of a new administrative housing regulation introduced in the beginning of 2002. Also, the number of purchases declined in the beginning of 2002.

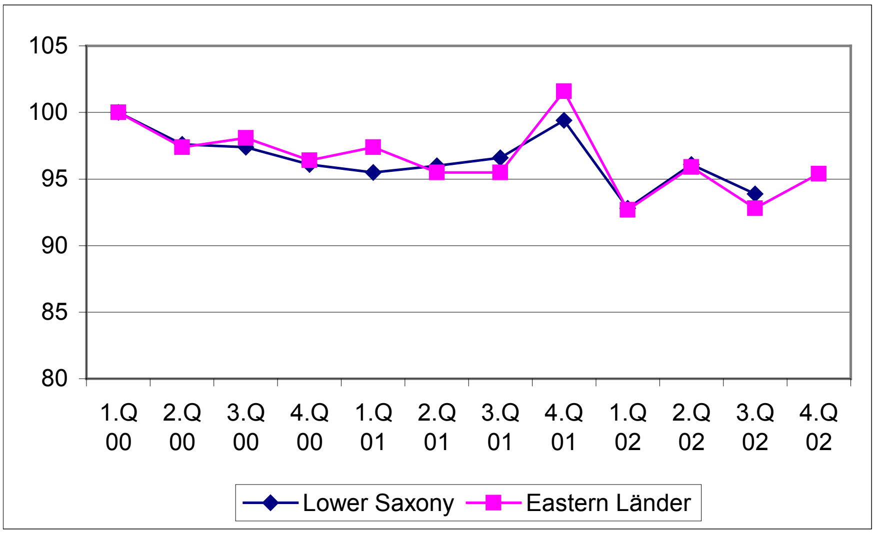


Figure 7: Price indices for housing in Germany

Overall, in spite of some difficulties, the main finding is that it seems feasible to compile a hedonic price index for owner occupied housing in Germany at least on a quarterly basis.

Session 4: Quality Control of Price Index Production Processes

Moderator: George Beelen, Statistics Canada

Summary of Session

The papers considered under this topic covered quality measurement and evaluation, outlier detection for editing, and quality assurance of price index production processes.

Estimates of sampling errors are important in allocating sample sizes, in production as a tool for editing, and for users of statistics in judging quality. The use of two-dimensional designs with elements of non-probability sampling methods, annually updated samples and complex estimators creates problems when estimating sampling variances for CPI-statistics.

Anders Norberg's paper considers variation in price changes by use of analysis of variance models with one locality and one item effect and an error corresponding to the selection of product offers. The implications for sample allocation are discussed.

Variance estimators in analytical forms, expressed in schemes of resampling procedures and model-based formulas are tested in a large simulation study. A 'random groups' method is proposed. This method makes it possible to estimate the variance for complex functions of index links such as the annual change of quarterly average inflation rates.

In discussion of the paper, Bert Balk observed that there are three possible approaches to estimating price index sampling error: normal statistical methods in cases of probabilistic sampling; stochastic approaches using analysis of variance and empirical approaches. He observed that empirical approaches appear to result in higher estimated sampling error than the other two approaches. Several comments noted the importance of more research on the issue of confidence intervals for price indices using non-probabilistic samples.

Hedonic quality adjustment methods are relatively expensive. The paper by Fenwick et al discusses strategic and practical issues in the choice of explicit quality adjustment methods, particularly the use of hedonic regression. The paper describes an approach to choosing *a priori* those components of the CPI where hedonic quality adjustment methods can be applied to maximum benefit to the statistical integrity of the index, thereby providing a means of strategically focusing resources in the application of hedonics.

In discussion of the paper, David Fenwick estimated that hedonic quality adjustment methods were potentially justified for only a relatively small number of CPI components, perhaps 10-20 products groups, based on the approach used. Y. Finkel suggested three broad categories of products as likely candidates for the use of hedonic quality adjustment based on importance and degree of price change criteria: seasonal products, technology products and owner-occupied housing.

Another important activity affecting quality is statistical editing. It is important to focus editing resources on those prices most likely in need of adjustment. A note by Carsten Boldsen Hansen on the use of the Hidiroglou-Berthelot method in the Danish CPI was presented. The HB method identifies an acceptance interval based on median and quartiles of price changes where those price changes outside the interval are identified as outliers and therefore as possible errors. The note described a step by step procedure used to apply the method in the Danish CPI. For the Danish CPI all outliers influencing the monthly elementary aggregate indices by more than one percent are checked, as well as extreme outliers even if they have no significant influence. Based on experience to date, the HB method as applied seems to identify too many outliers, a very high proportion of which turn out to be correct. Some adjustments in the parameters used are being considered to reduce the number of outliers detected.

In discussion of the paper, the assumption of a normal symmetric distribution of price relatives was questioned. Testing of this assumption seems warranted. The use of logs of price relatives may be appropriate. Sensitivity testing using different thresholds for outlier detection seems an obvious step, but it was noted that few statistical agencies' price index systems are designed to perform such testing cost-effectively.

Price index production processes are varied and complex, and are conducted under conditions of strict production deadlines and usually non-revisability. Thus quality assurance processes to ensure that established work procedures are followed and contribute to quality are crucial. Two papers in the session from Switzerland and Israel discussed quality management approaches used in the CPI context. The Swiss CPI paper discusses a Total Quality Management approach to the monthly CPI. The European Foundation for Quality Management (EFQM) approach was applied in the Israel CPI.

In discussion of these papers, it was observed that the use of the word "error" (as in having a goal of "zero error") in the context of quality assurance processes is in a different sense of the word than "statistical error". In TQM for example, "errors" are defined as instances where established standards and procedures have not been followed, whether or not that leads to error in the data. But statistical error refers to the uncertainty such sampling errors, which are inherent in any sampling approach. The use of Quality Management consultants was felt to be appropriate as advisors but not in deciding suitable procedures in the price index production context. Y. Finkel suggested that one effective strategy to increase documentation of processes was to it in formalized workplans and performance evaluations of staff, as is done in Israel.

Recommendations for Statistical Agencies

1. Statistical agencies should be encouraged to conduct more research into the production of confidence interval estimates for price indices, whether based on probability or non-probability sampling approaches. Traditional standard error estimates are more feasible in the context of probability sampling. Empirical or stochastic approaches to sample error estimation are possibilities when non-probability sampling is used.

2. The distribution of errors should be analysed when designing outlier detection strategies for statistical editing of price change. Statistical agencies should investigate means of making their processing more amenable to performing sensitivity testing of different statistical editing options.

Anders Norberg[164]: Comparison of Variance Estimators for the Consumer Price Index

Abstract Estimates of sampling errors are of importance for survey designers when allocating sample sizes, in production as a tool for output editing and for users of statistics in decision making. The use of two-dimensional designs with elements of non-probability sampling methods, annually updated samples and complex estimators creates problems when estimating sampling variances for CPI-statistics.

In this paper the character of variation in price changes is studied by use of analysis of variance models with one locality and one item effect and an error corresponding to the selection of product offers. This yields valuable information for the sample allocation work.

Variance estimators in analytical forms, expressed in schemes of resampling procedures and model based formulas are tested in a large simulation study. A random groups method is proposed. This method makes it possible to estimate the variance for complex functions of index links such as e.g. the annual change of quarterly average inflation rates.
.

1 Introduction

Measures of statistical errors can be used:

- to inform the users on the quality of the statistics,
- in the choice of statistical parameters,
- as a basis for a better allocation of resources in the production process and
- to detect possible serious errors in the data (output editing).

The study in question here was initiated to meet the needs of the first of the above-mentioned objectives for the users of the Swedish Consumer Price Index, KPI. One of the methodology approaches that were tried is, however, particularly appropriate for resource allocation.

Dalén and Ohlsson (1995) worked out a variance expression for the KPI's short-term index link in the daily necessities system. They also put forward a proposal for an estimator for the variance. The daily necessities system is a two-dimensional sample. Here we have one probability sample of outlets selling daily necessities and three samples of products, one for each of the three outlet chains that have a dominant market share.

For goods and services other than daily necessities, the price collector makes a choice among the products available within a more or less broad definition of the representative item. The person collecting the prices also decides which replacement products should be chosen when he or she must find a replacement product. This person should also make an evaluation on the quality of the products. On a micro level, "changes in price" therefore very often occur when replacement products are chosen. The choice of product offers, including quality evaluations, may influence the index and its variance more than the choice of outlet or representative item. This has not been sufficiently investigated in Sweden. It could even be that the "sample" of price collectors is significant for the variation in data. We can expect the variance estimator suggested by Dalén and Ohlsson (1995) to overestimate variance if it is applied to the sample dimensions outlets and representative

[164] Statistics Sweden

items, and if the variance between products is large in comparison to the variance between the representative items.

In this paper, a further number of proposals for variance estimators are given. Simulations have been carried out to compare them. These simulations have been carried out partly on completely synthetic populations and partly on populations taken from KPI data.

Balk (1989) and (1991), Caron and Duval (1994) and Shoemaker and Johnson are examples of work done on estimation variances for price indices.

Bases, formulae and calculation systems in Sweden's consumer price index today are described in section 2. Section 3 gives details of eight variance estimators. The results of the simulations with these estimators are reported in section 4. Certain applications and results are presented in section 5. The conclusions of the study, with recommendations for methods, are given in section 6.

2 General principles and methodology of the Swedish KPI today

The basic principle

The 1943 Index Commission (SOU, 1943) undertook a review of the foundations of the index. The idea that, as a matter of principle, the index should refer to the same standard of living in two different time periods was explicitly outlined. The method of calculating the index, still termed the "Cost of Living Index", was revised. From then on, it was computed as a chained index with annual links, in which the weights apply to the current year. The basic link in the index is defined as the price change between two successive Decembers, with quantity weights representing the whole calendar year between them. Longer-term comparisons could be obtained by multiplying successive links together into a chained index.

The chained index today

The current index base year is 1980. For each successive new link, weights are recalculated based on new information. We make a distinction between this long-term link (*L*), which uses quantity weights Q_y from year y and the short-term link (*K*), which uses quantity weights Q_{y-1} from year *y-1*. The definitions of the links are

$$L_{y-1,12}^{y,12} = \frac{\sum_k P_k^{y,12} Q_k^y}{\sum_k P_k^{y-1,12} Q_k^y} \quad \text{and} \quad K_{y-1,12}^{y,m} = \frac{\sum_k P_k^{y,m} Q_k^{y-1}}{\sum_k P_k^{y-1,12} Q_k^{y-1}} \tag{1}$$

where summation is over products with subscripts *k*.

The chained index from base year *0* to month *m* in year *Y* will now be[165]:

$$KPI_0^{Y,m} = I_0^{0,12} \cdot L_{0,12}^{1,12} \cdot \ldots \cdot L_{y-1,12}^{y,12} \cdot \ldots \cdot L_{Y-2,12}^{Y-1,12} \cdot K_{Y-1,12}^{Y,m} \tag{2}$$

[165] The extra link from 0 to 0.12 is needed in order to use a full year as the index reference period. Its exact definition is: $I_0^{0,12} = L_{-1,12}^{0,12} \Big/ \frac{1}{12} \sum_m K_{-1,12}^{0,m}$

12-month changes

When 12-month changes are published, the long-term index component is replaced by the corresponding short-term component. One reason for this procedure is that the long-term component includes substitution effects that do not represent the current 12-month period and another reason is the need for international comparability. The price change from *Y-1,m* to *Y,m* is thus calculated as:

$$Infl_{Y-1,m}^{Y,m} = K_{Y-1,12}^{Y,m}\, K_{Y-2,12}^{Y-1,12} \big/ K_{Y-2,12}^{Y-1,m} = \frac{KPI_0^{Y,m}}{KPI_0^{Y-1,m}} \frac{K_{Y-2,12}^{Y-1,12}}{L_{Y-2,12}^{Y-1,12}} \qquad (3)$$

The 12-month changes according to (3) are often used for monitoring inflation. For example, the inflation target of the Swedish Riksbank (the Central Bank) is the moving average of twelve 12-month changes.

Index aggregation at the higher level

The National Accounts provide values of consumption, *V,* at the higher KPI aggregation levels. These are defined as "price times quantity", in our notation, $V=P*Q$. Up-to-date National Accounts consumption values today exist for more than 100 consumption categories. During the annual weight revision new National Accounts values are brought into the index. Below the level where National Accounts values are available, it is necessary to vary procedures. Household Budget Surveys (HBS) are used for breaking down many NA categories into smaller groups. For food, Statistics Sweden aggregates scanner data received from the three leading chains in the retail market. As December is the linking month, the item weights are computed proportional to these annual consumer values price-back-dated and price-up-dated for the long-term and short-term links respectively.

Index aggregation at lower levels

At the lowest level, elementary aggregation, the KPI generally uses the RA formula in the sense of Dalén (1992).

$$I_{y-1,12}^{y,m} = \frac{\sum_k w_k p_k^{y,m} /(p_k^{y-1,12} + p_k^{y,m})}{\sum_k w_k p_k^{y-1,12} /(p_k^{y-1,12} + p_k^{y,m})} \qquad (4)$$

New index construction

In the Swedish KPI publication for January 2005 and onwards, the KPI numbers will be computed by using an improved index construction. Like now, the KPI will also subsequently be computed as a chain index with annual links. In the new construction the annual link will measure how much the average price level in the year concerned has changed from the average price level of the preceding year. The CPI basket of the annual link will reflect a blend of consumption patterns of the year concerned and the preceding year, according to a Walsh-formulae.

The new index construction will also make it possible to use more data from the national accounts and to use them more accurately for the index weights. A further change is that sub-indices on the lowest levels of aggregation will be computed as geometric means of price relatives.

Central and local price collection

There are basically two different modes of price collection. For most services and some goods, the central staff collects the prices, either by telephone or using a small-scale postal survey. This procedure is referred to as *central price collection*. Owner occupied housing and rents are also parts of the central price collection. Alcohol and medicine are sold by state monopolies in Sweden and price indices are computed without sampling errors.

For many goods and services, there is *local price collection*. The specifications for so called representative items are established centrally. In some cases, the specifications are broad and the price collectors can select the "most sold" product (a product offer) within the specification in a selected outlet. In other cases, the specifications are tight and the price collectors pick corresponding products in the sampled outlet or, if one is not found, omit the product.

Prices are collected by price collectors by visiting the outlets directly or by a telephone call. The collection process takes place on an optional day in the week in which the 15^{th} of the month occurs. In each outlet, from one up to as many as 500 prices are observed. In all, some 25 000 prices in 900 outlets are observed in the local price collection.

Sampling methods in local price collection

Sampling of representative items in the daily necessities system

For foodstuffs (except for fresh food, such as meat and vegetables) and other daily necessities, product sampling is carried out by Pareto π ps (see Rosén(2000)), using sampling frames provided by the three major retail chains in Sweden. These are estimated to be some 80% of all goods sold in supermarkets. The sampling frame covers products at a detailed level, where a unique price normally exists, for example " EAN = 7331040056126 Coca Cola Light, plastic bottle, 1.5 litre".

Three different product samples of 400 products each are created, one for each of three outlet chains. The product samples are then matched to the outlet sample according to the chain to which a sampled outlet belongs. Only product offers in the sampled outlet are thus included. This reduces the effective product sample size in each outlet to some 250-300 product offers. There are 40 supermarkets and 9 hypermarkets in the outlet sample.

Sampling of representative items in clothing and other local prices systems

For clothing, furniture, other goods sold in the retail trade and services, such as restaurants, the representative items are chosen and specified judgmentally, in a manner typical for consumer price index methods in most countries.

Outlet sampling in local price collection

In product groups, where local price collection is used, outlets are divided into 50 retail trade and service strata according to SNI code (Swedish Standard Industrial Classification, which closely follows NACE, Rev. 1, the EU standard). In each stratum, a sample of outlets is drawn from the Business Register by an order π ps technique. This first gross sample is drawn about 6 months before the year in which the sample is to be used. This sample is then screened, both in the central office and by the price collectors visiting the outlets, and some of the sampling units initially drawn are excluded for various reasons. For example, they may be offices rather than outlets, or they may not sell any products fitting the sampled items.

A sample design with overlapping panels is implemented by the use of random numbers permanently associated with every outlet in the sampling frame. Sampling rotation is performed so that 20% of the random numbers are changed every year. Combined with changes in the sampling frames, this results in some 70-75% of outlets remaining in the sample from one year to the next. See Ohlsson (1990 and 1995) or Statistics Sweden (2001) for a description of this technique.

New and disappearing products and outlets / replacements

New samples are introduced when a new index link is started up in December, when both the old and the new item and outlet samples are measured. In this way, an *overlap* is created, so that the old sample is used for back comparisons and the new sample for forward comparisons, without any explicit quality adjustment. To the extent that the market is in equilibrium, so that the price differentials between the old and the new sample in December reflect genuine consumer valuation of quality differences, and that both samples adequately represent the population of product offers, the estimator of price changes is unbiased. For clothing and footwear, the former requirement is not fulfilled and an adjustment for obsolescence of the old sample is made. Where an outlet remains in the sample from one year to the next, the sampled product offers are normally not exchanged in December.

The *replacement* of a particular product offer with another in the same outlet is caused by the disappearance or reduced significance of a product offer. In this case, a quality adjustment is normally done.

Outlets are not replaced during the year. New outlets are only introduced in the course of updating samples in December. In those, very few, cases where an outlet is closed down or price measurements cease to be possible for some other reason, that outlet's products offers are deleted.

Quality change

Quality adjustment in the clothing system

The rapid changes of items in the clothing market make advanced and complex methods for dealing with replacements and quality adjustments necessary. For this reason, *hedonic models* of the relationship between clothing prices and characteristics are formulated and estimated. The adjustment for obsolescence is also significant. Due to the needs of the hedonic model method, data collection and processing is more expensive than for other product areas.

Quality adjustment for other products in the local price collection

Here, the price collectors perform the adjustments. By quality difference is understood a difference in function, comfort, durability, security, guarantees and easiness of handling etc.. Differences in quality are to be valued from the viewpoint of the consumer. The price collector should try to assess how the average consumer experiences differences. This is difficult and in practice it means that he/she will have to use his/her own assessment of the differences.

Non-quality adjustment products in local price collection

For products where quality adjustments are not made, but package size has been altered, only new package sizes where the quantity change is less than 50% are accepted as replacements. A proportional adjustment is then made so that the price effectively becomes a price per quantity unit.

In other cases, where a product offer can no longer be found, it is deleted and the price change is imputed from the rest of the product offers in the product group.

3 Variance estimation methods

Variance is a statistical measure of variation for a random variable. When, as is the case here, the random variable is a measurement of a studied phenomenon (inflation) we can interpret the variance as a measure of statistical uncertainty. The uncertainty we are analysing here is restricted to the fact that inflation is calculated using samples of outlets, representative items and product offers - it is not possible to collect data on all transactions in society. There are other sources of errors in statistics, which can be considered more serious. In Dalén (1999), sources for the bias and size of them are assessed.

Problems

- One condition for the concepts of variance and variance estimates is that we, in principle, have statistical samples, i.e. samples that have been selected using random methods and with known probability. This is not always (read: seldom) the case in CPI-surveys.
- Two-dimensional sampling of outlets and products is not covered in the traditional statistical literature. This design is the only design for probability sampling that is practical to handle for a CPI-survey.
- Outlet samples are rotated once a year in a coordinated way.
- The estimators are complex, including index links from more than one year.

For the Swedish KPI the condition on statistical samples is fulfilled with regard to the outlet sample and the product sample in the daily necessities system. For clothes and other locally priced items, we must compromise with the principles and consider the sample of items and/or product offers as random. We can either see them as:

- A random sample of representative items in the first stage and the collection of the "most sold" product offer within the representative item, which would not mean a choice in the second stage, neither deliberate nor random.
- A random sample of representative items in the first stage and a random (at least arbitrary) sample of product offers in a second stage, independently drawn in each outlet.
- A stratification of the markets by the representative items, which therefore covers all or a large part of the market, and a random (arbitrary) sample of product offers in one single stage, independent between outlets.

The inclusion probability

The inclusion probabilities for outlets and representative items and product offers are essential for the estimation of variance. For judgemental samples we have to impute values. There are different situations. For clothing the sample of representative items cover a large part of the market and the “inclusion probabilities” are high but there are many product offers to chose among and the proportion is low. For veterinarian services the CPI-statistician has chosen one specific service among many and there is perhaps no more than one product offer to chose among (compare the Swedish daily necessities system).

Taylor linearisation

Dalén and Ohlsson (1995) have derived the variance expression for an index based on a two-dimensional sample and suggest an estimation of the variance. The variances and variance estimators have three terms that can be interpreted as variance between products, between outlets and an interaction term.

Let m_g be the sampled number of products in product group *g*, n_h be the number of outlets in outlet group *h* and v_{gh} be the weighting, based on sales during one year, for the combination (*g, h*). The weighting is standardized so that $\sum_g^G \sum_h^H v_{gh} = 1$.

π_{gi}^R and π_{hj}^C are sample probabilities for product offer *i* within product group *g* and outlet *j* within outlet group *h*. Denote the price of product offer *i* within product group *g* and outlet *j* as p_{ij}.

Let:

$$I_{ij} = \begin{cases} 1 \text{ if product offer } i \text{ is available in outlet } j \\ 0 \text{ otherwise} \end{cases}$$

$$f_{ij}^0 = \frac{p_{ij}^0}{(p_{ij}^0 + p_{ij}^1)/2} \qquad f_{ij}^1 = \frac{p_{ij}^1}{(p_{ij}^0 + p_{ij}^1)/2} \tag{5}$$

For each *i*, w_i^R is a weighting for product offer *i* and for each *j*, w_j^C is a weighting for outlet *j*. These weightings depend on the size (turnover) of the particular sample object and sample procedure (OSU, PPS, etc.) The weighting is standardized so that $\sum_{i \in U_g^R} w_i^R = \sum_{j \in U_h^C} w_j^C = 1$ for each g and h.

Let

$$\hat{X}_{gh} = \sum_{i \in U_g^R} \sum_{j \in U_h^C} 1_{ij} w_i^R w_j^C f_{ij}^0, \qquad \hat{Y}_{gh} = \sum_{i \in U_g^R} \sum_{j \in U_h^C} I_{ij} w_i^R w_j^C f_{ij}^1 \tag{6}$$

$\hat{X}_{gh} = 0$ where product *i* is not traded in outlet *j*, the same applies to $\hat{Y}_{gh}$

$$\hat{I}_{gh} = \frac{\hat{Y}_{gh}}{\hat{X}_{gh}} \text{ is a "cell index".} \tag{7}$$

The KPI's short-term index link becomes $\hat{I} = \sum_g^G \sum_h^H v_{gh} \hat{I}_{gh}$ (8)

Previously $\hat{e}_{ij}^{gh} = 1_{ij}(f_{ij}^1 - \hat{I}_{gh} f_{ij}^0)$, $\hat{e}_{i.}^{gh} = \frac{1}{m_g} \sum_{i \in S_g^R} \hat{e}_{ij}^{gh}$, $\hat{e}_{.j}^{gh} = \frac{1}{n_h} \sum_{j \in S_h^C} \hat{e}_{ij}^{gh}$ (9)

Dalén and Ohlsson suggest, after some the following variance estimator for $\hat{I}$.

$$\hat{V}_{D\&O} = \hat{V}_{PRO} + \hat{V}_{BUT} + \hat{V}_{INT} \tag{10}$$

$$\hat{V}_{PRO} = \sum_g \frac{1}{m_g(m_g - 1)} \sum_{i \in S_g^R} \left(1 - \pi_{gi}^R\right) \left\{ \sum_h \frac{v_{gh}}{\hat{X}_{gh}} \hat{e}_{i.}^{gh} \right\}^2 \tag{11}$$

$$\hat{V}_{BUT} = \sum_h \frac{1}{n_h(n_h - 1)} \sum_{j \in S_h^C} \left(1 - \pi_{hj}^C\right) \left\{ \sum_g \frac{v_{gh}}{\hat{X}_{gh}} \hat{e}_{.j}^{gh} \right\}^2 \tag{12}$$

$$\hat{V}_{INT} = \sum_g \sum_h \frac{v_{gh}^2}{\hat{X}_{gh}^2} \frac{1}{n_h(n_h - 1)} \frac{1}{m_g(m_g - 1)} \cdot \left\{ \sum_{i \in S_g^R} \sum_{j \in S_h^C} \left(1 - \pi_{hj}^C\right)\left(1 - \pi_{gi}^R\right)\left(\hat{e}_{ij}^{gh} - \hat{\bar{e}}_{i.}^{gh} - \hat{\bar{e}}_{.j}^{gh}\right)^2 \right\} \quad (13)$$

Standard formulae for one- and two-stage OSU

Under the assumption that the setting of prices in a market, say retail trade of furniture, is not decided by one outlet changing their price strategy or by the price for one type of furniture (a representative item) changing overall, it could be a reasonable approximation to consider the collected data as generated as a simple random sample, SRS, from the whole population of products offers in the country. Variance estimators according to standard formulae for stratified SRS from the whole market, is a reference method.
Two-stage sampling, with stratified SRS of outlets in the first stage and independent stratified SRS of products offers per outlet, and vice versa, are also interesting and easily calculated alternatives.

Replication techniques

A number of techniques based on replication have been suggested for variance estimation in the last 50 years. Independent and dependent random groups, balanced half-samples, the jackknife and the bootstrap are well-known examples (Wolter (1985)). In certain situations, one or more of them are not suitable.

IN THE CASE OF THE KPI, THERE ARE SOME CONSIDERATIONS TO BE MADE:

- There are many strata, both in the outlet and in the product dimension. The sample sizes in some outlet strata is as small as six, and, in the product dimension, only two or even one representative item.
- Some sample units, outlets as well as representative items, are selected with certainty and some are selected with high inclusion probabilities, while most sample units are selected with small inclusion probabilities.
- The samples are updated once a year. We require a variance estimating procedure that, in the best possible way, measures the effect of rotating outlet samples, the introduction of new goods and services and new item weights on sampling variance.

We propose a repetitive use of the dependent random groups method with the minimum number of random groups, k=2, in each repetition.

Repeated random groups (1)

The basics are this: An estimator $\hat{Y}$ is based on a sample. The full sample is divided into *k* sub-samples, each with *1/k* of the sample units, so that the design of each sub-sample in best manner is the same as the full sample regarding stratification, rotation, etc. The estimator $\hat{Y}_\alpha$, based on sub-sample α looks like $\hat{Y}$. An estimator of the variance of $\hat{Y}$ is $\frac{1}{k-1}\left(\hat{Y} - \hat{Y}_\alpha\right)^2$. The average of all *k* variance estimators is proposed:

$$\hat{V}\left(\hat{Y}\right) = \frac{1}{k} \sum_{\alpha=1}^{k} \frac{1}{k-1} \left(\hat{Y} - \hat{Y}_\alpha\right)^2 . \quad (14)$$

The value $k = 2$ is an extreme, here a sub-sample contains of half of all data and we create the estimator $\hat{Y}_{½}$. An estimator of the variance for $\hat{Y}$ is $\left(\hat{Y} - \hat{Y}_{½}\right)^2$. This estimator has only one degree of freedom and we cannot extract much more information from the two half-samples. Now we estimate the variance many times by repeatedly drawing "half samples" and take the mean value of estimated variances to ensure sufficient quality in the final variance estimator.

Up to this point, we have used established theories. How can a good variance estimator be produced when we meet a two-dimensional sample? The following procedure is a trial, which has been tested in a simulation study.

Let us draw a "half sample" in both dimensions so that, as a result, we only include a quarter of all data. Study this random breakdown of a total sample:

x_{11}	x_{12}
x_{21}	x_{22}

Let x_{11}, x_{12}, x_{21} and x_{22} be the sums of a variable X for the same number of objects in the four boxes. With a little algebra, we get the following

$$\frac{1}{4}\sum_{r=1}^{2}\sum_{c=1}^{2}(4 \cdot x_{rc} - \sum_{r=1}^{2}\sum_{c=1}^{2} x_{rc})^2 =$$

$$= \frac{1}{2}\sum_{r=1}^{2}(2 \cdot (x_{r1} + x_{r2}) - \sum_{r=1}^{2}\sum_{c=1}^{2} x_{rc})^2 + \frac{1}{2}\sum_{c=1}^{2}(2 \cdot (x_{1c} + x_{2c}) - \sum_{r=1}^{2}\sum_{c=1}^{2} x_{rc})^2 +$$

$$+ \frac{1}{2}\left[(2 \cdot (x_{11} + x_{22}) - \sum_{r=1}^{2}\sum_{c=1}^{2} x_{rc})^2 + (2 \cdot (x_{12} + x_{21}) - \sum_{r=1}^{2}\sum_{c=1}^{2} x_{rc})^2\right] \qquad (15)$$

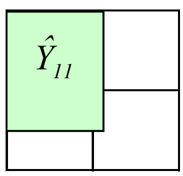
We can call these three terms between-outlet, between-items and interaction sums of squares

The finite population corrections for certain large sample probabilities can be significant for a consumer price index. The use of explicit corrections is impractical, particularly as there are many strata and the sample probabilities for outlets and products vary and are different from year to year. Denoting the inclusion probability with π, the finite correction factor is $(1 - \pi)$.

Let us test the following procedure. We are looking for the individual value of k ("number of random groups", which no longer need to be a whole number) per sampled object (individual outlet and product respectively) which is such that $(1 - \pi) \cdot \frac{1}{k - 1}$ is equal to 1.

In this way, the effect of the finite population correction is hopefully included in the procedure to select a replica of the sample. We choose to set this to 1 because when k=2 and no finite sampling correction factor is considered, the expression in front of the squared deviation is 1. In this way we get a very simple calculation formula. A little algebra leads to the "half sample" probabilities $p_i = 1/(2 - \pi_i)$ that are in the range 0.5 – 1.0.

Because we now select slightly over half of all outlets and items/products in a "half sample", it is in practice appropriate to select only one of the four "half samples", say the one based on the box $r=1$ and $c=1$.



With Q repetitions of the procedure, we get the variance estimator

$$\hat{V}_{RG(1)}\left(\hat{Y}\right)=\frac{1}{Q}\sum_{q}^{Q}\left(\hat{Y}_{11q}-\hat{Y}\right)^{2} \quad (16)$$

Random groups estimation with two-stage sampling (2) and (3)

According to Wolter (1985), page 31, when selecting a multi-stage sample, random groups should be created by, in a random way, dividing the sample of primary selected units, PSU, into *k* groups, while keeping the original sample design as much as possible. All samples in later stages should remain undivided. Also, since finite population corrections are not used, variances may be underestimated.

Assume we measure variable *Y* in a two-dimensional population. The measurement value is y_{ij} for the *i* th sample unit in the first dimension and the *j* th sample unit in the second dimension.
Let $\hat{Y}_i$ be an estimator of Y_i based on any design with a probability sample of second-stage units. An estimator of a total *Y* from a simple random sample of *n* PSUs, with the replacement from a population with *N* is

$$\hat{Y}=\frac{N}{n}\cdot\sum_{i=1}^{n}\hat{Y}_i \quad (17)$$

The usual variance estimator for the HT-estimator of a total *Y* from a SRS without replacement of *n* PSUs from a population with *N* and SRS without the replacement of m_i second-stage units from M_i is

$$\hat{V}_1(\hat{Y})=\frac{N^2}{n}\cdot\left(1-\frac{n}{N}\right)\cdot s_b^2+\frac{N}{n}\cdot\sum_{i=1}^{n}M_i^2\left(\frac{1}{m_i}-\frac{1}{M_i}\right)s_{wi}^2 \quad (18)$$

where $s_{wi}^2=\frac{1}{m_i-1}\cdot\sum_{j=1}^{m_i}\left(y_{ij}-\bar{y}_i\right)^2$ (19)

Analysis of variance have shown that price changes in many markets occur in an apparently random manner; they cannot be explained by the effects of either the outlets or the items. Let us assume that we have a two-stage sample from a two-dimensional population with *NM* objects in which the measurement values y_{ij} have been generated as random figures, $i=1,...,N$ and $j=1,...,M$. On "average" therefore, all s_{wi}^2 are the same and can be taken as the mean of the measurement values for all first-stage units according to the following:

$$s_{wi}^2=\frac{1}{m-1}\sum_{j}^{m}\left(y_{ij}-\bar{y}_i\right)^2 \hat{=} n^2\cdot\frac{1}{m-1}\sum_{j}^{m}\left(\frac{1}{n}\sum_{i=1}^{n}y_{ij}-\frac{1}{n}\sum_{i=1}^{n}\bar{y}_i\right)^2=n^2\cdot s_{\bar{w}}^2 \quad (20)$$

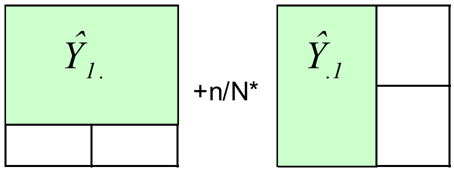
We can now propose the following variance estimator:

$$\hat{V}_2\left(\hat{\bar{Y}}\right)=\frac{1}{N^2\cdot M^2}\left[\frac{N^2}{n}\cdot\left(1-\frac{n}{N}\right)\cdot s_b^2+\frac{N}{n}\cdot\frac{M^2}{m}\cdot\left(1-\frac{m}{M}\right)\cdot n^2\cdot s_{\bar{w}}^2\right]$$

$$=\frac{1}{n}\cdot\left(1-\frac{n}{N}\right)\cdot s_{\bar{b}}^2+\frac{n}{N}\cdot\frac{1}{m}\cdot\left(1-\frac{m}{M}\right)\cdot s_{\bar{w}}^2 \quad (21)$$

where $s_{\bar{b}}^2=\frac{1}{n-1}\cdot\sum_{i=1}^{n}\left(\bar{y}_i-\frac{1}{n}\sum_{i=1}^{n}\bar{y}_i\right)^2$ (22)

With a random groups design with a total variance that is the sum of the between-outlets, between-products and interaction variances, we now form a variance estimator consisting of the between-outlets variance plus the sampling fraction for the outlet sample (n/N) multiplied by the between-products variance, or vice versa[166].

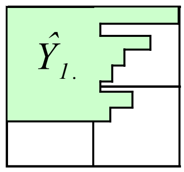


$$\hat{V}_{RG(2)}\left(\hat{Y}\right)=\frac{1}{Q}\sum_{q=1}^{Q}\left[\left(\hat{Y}_{11+12}-\hat{Y}\right)^{2}+\frac{n}{N}\left(\hat{Y}_{12+22}-\hat{Y}\right)^{2}\right] \tag{23}$$

We convert this variance estimator into one combined procedure for repeated random half-sampling. Say we have already chosen a "half sample" of first-stage units (outlets[2]) with the probabilities $p_{(1)}=1/(2-\pi_{(1)})$ and all second-stage units (items). $\pi_{(1)}=n/N$ for the first stage units and $\pi_{(2)}=m/M$ for the second-stage units. Now we would also select a "half sample" of second-stage units (items) so that $\pi_{(1)}\cdot\left(1-\pi_{(2)}\right)\cdot\frac{1}{k-1}$ becomes 1, which means that the second-stage units (items) should be selected with the probabilities

$$p_{(2)}=\frac{1}{k}=\frac{1}{1+\pi_{(1)}\cdot\left(1-\pi_{(2)}\right)}. \tag{24}$$

This can be illustrated this way. Depending on the sampling probabilities in the first and second stages of sampling, for each first-stage unit we select varying proportions of second stage units into the half-sample.



We see, that if a $\pi_{(1)}$ is close to zero, we shall select all second-stage units within the PSUs in the text-book manner.

We see also, that if a $\pi_{(1)}=1$, we shall select second-stage units within the PSU in the same way as we have selected the PSUs when $\pi_{(1)}<1$.

The estimator based on Q replications is

$$\hat{V}_{RG(3)}\left(\hat{Y}\right)=\frac{1}{Q}\sum_{q}^{Q}\left(\hat{Y}_{11q}-\hat{Y}\right)^{2} \tag{25}$$

[166] Which is considered as first- and second-stage units of outlets and products depends on the homogeneity of the descriptions of the representative items.

TABLE 1 THE PROPORTIONS OF SECOND-STAGE (2) UNITS TO BE SELECTED INTO A “HALF-SAMPLE” PER FIRST-STAGE UNIT (1), DEPENDING ON THE SAMPLING PROBABILITIES IN THE FIRST AND SECOND SAMPLING STAGES

	$\pi_{(2)}$			
$\pi_{(1)}$	0.0	0.5	0.9	1.0
0.0	1.00	1.00	1.00	1.00
0.5	0.67	0.80	0.95	1.00
0.9	0.53	0.69	0.92	1.00
1.0	0.50	0.67	0.91	1.00

TABLE 2 THE OVERALL PROPORTIONS OF UNITS TO BE SELECTED INTO A “HALF-SAMPLE”, DEPENDING ON THE SAMPLING PROBABILITIES IN THE FIRST AND SECOND SAMPLING STAGES

	$\pi_{(2)}$			
$\pi_{(1)}$	0.0	0.5	0.9	1.0
0.0	0.50	0.50	0.50	0.50
0.5	0.44	0.53	0.63	0.67
0.9	0.48	0.63	0.83	0.91
1.0	0.50	0.67	0.91	1.00

Table 3 presents the testing of this procedure for variance estimation for eight synthetic populations. A model generates PSUs with a large within-PSU variance. We compare the variance estimator $\hat{V}_{RG(3)}\left(\hat{Y}\right)$ with a "half sample" estimator of variance for PSUs only, expecting an underestimation for the latter. The populations have 100 PSUs and 100 second-stage units for each PSU. Table 3 shows that the variance estimator without "within PSU component" has a negative bias and also that our proposed estimator $\hat{V}_{RG(3)}\left(\hat{Y}\right)$ is positively biased, but generally not as much.

Table 3 Estimated variances after two-stage sampling with the repeated random group method in relation to an empirical true variance. The result of simulations with varying sample probabilities from populations in which the measurement variable is a linear function of a row and a column effect.
Simple = random groups defined only for PSU,
Double = random groups defined for both first- and second-stage units = $\hat{V}_{RG(3)}(\hat{Y})$.

	$\pi_{(2)}$							
	0.0–0.4		0.0–1.0		0.6–1.0		1.0	
$\pi_{(1)}$	Simple	Double	Simple	Double	Simple	Double	Simple	Double
0.0–0.4	0.89	1.12			1.00	1.07		
0.0–1.0			0.88	1.10				
0.6–1.0	0.60	1.12			0.73	1.10	1.01	1.01
0.8–1.0			0.59	1.43				
1.0			0.00	1.65				

Footnote: In the categories 0.0-0.4 and 0.0-1.0, the sample probabilities between 0.0 and 0.1 are replaced by 0.1. The categories 0.0-1.0, 0.6-1.0 and 0.8-1.0 contain a number of units which have been definitely selected.

Model-based estimation

Professor Daniel Thorburn, University of Stockholm and Statistics Sweden has contributed with a theory we call model-based variance estimation. Dr Stefan Svanberg, Statistics Sweden, has contributed with a derivation of the calculation formulae for unbalanced situations. Thorburn and Svanberg are not responsible for possible misinterpretations or errors that have been made in the following.

SUPPOSE THAT, AT EACH POINT OF TIME, THE PRICES FOR A PRODUCT OFFER IS SET ACCORDING TO A STOCHASTIC PROCEDURE WITH FOUR RANDOM NUMBER GENERATORS. LET Y_{ijk} = THE LOGARITHM OF PRICE CHANGE FROM DECEMBER OF YEAR *Y-1* TO MONTH *M* OF YEAR *Y* FOR PRODUCT OFFER *K* IN OUTLET *J* FOR REPRESENTATIVE ITEM *I*. WE ASSUME THAT OBSERVATIONS ARE GENERATED ACCORDING TO

$$Y_{ijk} = \mu + \beta_i + \gamma_j + \delta_{ij} + \varepsilon_{ijk}, \qquad (26)$$

where $i = 1,..,m_g$, $j = 1,..,n_h$, $k = 1,..,r_{ij}$

μ is a general mean value without variance

β_i is a representative item effect, $i = 1,..,m_g$

γ_j is an outlet effect, $j = 1,..,n_h$

δ_{ij} is an interaction effect between a representative item and an outlet

ε_{ijk} is a product offer effect, $k = 1,..,r_{ij}$

Assume that β_i is normally distributed with the expected value 0 and variance σ^2_β, that γ_j is normally distributed with the expected value 0 and variance σ^2_γ, that δ_{ij} is normally distributed with the expected value 0 and variance σ^2_δ and finally that ε_{ijk} is normally distributed with the expected value 0 and variance σ^2_ε. In this case, Y_{ijk} is normally distributed with variance $\sigma^2_\beta + \sigma^2_\gamma + \sigma^2_\delta + \sigma^2_\varepsilon$.

Assume that we have a sample. Let $Y_{...}=\Sigma_{ijk} Y_{ijk}$, $Y_{i..}=\Sigma_{jk} Y_{ijk}$, $Y_{.j.}=\Sigma_{ik} Y_{ijk}$, $Y_{..k}=\Sigma_{ij} Y_{ijk}$ and $Y_{ij.}=\Sigma_k Y_{ijk}$.

In an analysis of variance, the sum of squares $SS_Y = \sum\sum\sum (Y_{ijk} - Y_{...})^2$ is split into the sums of squares which is explained by the factors

$$\sum_{i=1}^{m_g}(Y_{i..} - Y_{...})^2 + \sum_{j=1}^{n_h}(Y_{.j.} - Y_{...})^2 + \sum_{i=1}^{m_g}\sum_{j=1}^{n_h}(Y_{ij.} - Y_{i..} - Y_{.j.} + Y_{...})^2 \qquad (27)$$

and the residual sum of squares

$$SSE = \sum_{i=1}^{m_g}\sum_{j=1}^{n_h}\sum_{k=1}^{r_{ij}}(Y_{ijk} - Y_{ij.})^2 \qquad (28)$$

The SAS system has a procedure for analysis of variance with random effects, which allows a varying number of observations per cell, r_{ij}. The result provides the basis for the calculation of estimates of $\hat{\sigma}^2_\beta$, $\hat{\sigma}^2_\gamma$, $\hat{\sigma}^2_\delta$ and $\hat{\sigma}^2_\varepsilon$

```
proc glm
   data=pop;
   class outlet item;
   model pkvot=outlet item outlet*item;
   random outlet item outlet*item;;
run; quit;
```

When we have these estimators, we can also estimate variances for the mean values of logarithmic price ratios for a sample of representative items (m_g), outlets (n_h) and total number of product offers (r_{ij}), at least if the sample of product offers is fairly similar in size for all representative items and outlets. As long as we can generate data for a process, we get:

$$
\begin{aligned}
V(\sum_{i=1}^{m_g}\sum_{j=1}^{n_h}\sum_{k=1}^{r_{ij}} Y_{ijk}) &= V(\sum_{i=1}^{m_g}\sum_{j=1}^{n_h}\sum_{k=1}^{r_{ij}}(\beta_i+\gamma_j+\delta_{ij}+\varepsilon_{ijk})) = \\
&= V(\sum_{i=1}^{m_g}\sum_{j=1}^{n_h}\sum_{k=1}^{r_{ij}}\beta_i)+V(\sum_{i=1}^{m_g}\sum_{j=1}^{n_h}\sum_{k=1}^{r_{ij}}\gamma_j)+V(\sum_{i=1}^{m_g}\sum_{j=1}^{n_h}\sum_{k=1}^{r_{ij}}\delta_{ij})+V(\sum_{i=1}^{m_g}\sum_{j=1}^{n_h}\sum_{k=1}^{r_{ij}}\varepsilon_{ijk}) = \\
&= \sum_{i=1}^{m_g}V(\beta_i\sum_{j=1}^{n_h}r_{ij})+\sum_{j=1}^{n_h}V(\gamma_j\sum_{i=1}^{m_g}r_{ij})+V(\sum_{i=1}^{m_g}\sum_{j=1}^{n_h}r_{ij}\delta_{ij})+V(\sum_{i=1}^{m_g}\sum_{j=1}^{n_h}\sum_{k=1}^{r_{ij}}\varepsilon_{ijk}) = \\
&= \sum_{i=1}^{m_g}V(\beta_i r_{i.})+\sum_{j=1}^{n_h}V(\gamma_j r_{.j})+V(\sum_{i=1}^{m_g}\sum_{j=1}^{n_h}r_{ij}\delta_{ij})+V(\sum_{i=1}^{m_g}\sum_{j=1}^{n_h}\sum_{k=1}^{r_{ij}}\varepsilon_{ijk}) = \\
&= \sigma_\beta^2\sum_{i=1}^{m_g}r_{i.}^2+\sigma_\gamma^2\sum_{i=1}^{m_g}r_{.j}^2+\sigma_\delta^2\sum_{i=1}^{m_g}\sum_{j=1}^{n_h}r_{ij}^2+\sigma_\varepsilon^2\sum_{i=1}^{m_g}\sum_{j=1}^{n_h}r_{ij}
\end{aligned}
\qquad (29)
$$

When we take a sample from a finite population of outlets, representative items and product offers, we should be able to use the following estimator of variance.

LET THE TOTAL NUMBER OF OBSERVATIONS BE $r=\sum_{i=1}^{m_g}\sum_{j=1}^{n_h}r_{ij}=\sum_{i=1}^{m_g}r_{i.}=\sum_{j=1}^{n_h}r_{.j}$

WE PROPOSE THIS AS A VARIANCE ESTIMATOR FOR KPI DATA, DISREGARDING THAT THE ASSUMPTION THAT β_i, γ_j, δ_{ij} AND ε_{ijk} ARE NORMALLY DISTRIBUTED OBVIOUSLY IS NOT FULFILLED.

$$
\hat{V}\left[\frac{1}{r}\sum_i\sum_j\sum_k Y_{ijk}\right] = \left(1-\frac{m_g}{M_g}\right)\cdot\frac{\sum_{i=1}^{m_g}r_{i.}^2}{r^2}\cdot\hat{\sigma}_\beta^2+\left(1-\frac{n_h}{N_h}\right)\cdot\frac{\sum_{i=1}^{m_g}r_{.j}^2}{r^2}\cdot\hat{\sigma}_\gamma^2+ \\
+\left(1-\frac{m_g\cdot n_h}{M_g\cdot N_h}\right)\cdot\frac{\sum_{i=1}^{m_g}r_{.j}^2}{r^2}\cdot\hat{\sigma}_\delta^2+\left(1-\frac{r}{R}\right)\cdot\frac{1}{r}\cdot\hat{\sigma}_\varepsilon^2 \qquad (30)
$$

4 Simulations

The proposed variance estimators have been tested on a number of finite populations of outlets, representative items and products. We have created completely synthetic populations and populations with data in the KPI database. We have only one product stratum (g) and one outlet stratum (h). We consider stratum not to be a problem in this context. From each population, we have drawn thousands of two-dimensional samples with SRS in both dimensions. We consider varying sampling probabilities not to cause any special problem compared to sampling with equal probabilities. The samples are often large in relation to the populations (about 50 %) so that it has been necessary to pay attention to the finite population corrections.

In tests with clothing and furniture data, we have chosen two product offers per combination of representative item (i) and outlet (j).

The price index link has been calculated with the Swedish RA-formula in one step for the whole sample, as in the KPI.

Daily necessities

For daily necessities in the KPI there is a two-dimensional probability sample of outlets and products. First, the variance estimators are tested with extreme populations to see if any estimator fails completely under any conditions.

In population F, the between-items and between-outlets population variances are zero. When we take a sample of products and outlets, the estimated between-products and between-outlets variance terms are greater than zero. The structure of F is as follows:

	Outlet 1	Outlet 2	Outlet 3	Outlet 4	Outlet 5	Outlet 6
Product 1	+ x %	+ x %	+ x %	- x %	- x %	- x %
Product 2	+ x %	+ x %	+ x %	- x %	- x %	- x %
Product 3	- x %	- x %	- x %	+ x %	+ x %	+ x %
Product 4	- x %	- x %	- x %	+ x %	+ x %	+ x %

Table 4 Average variance estimates (*10^4) for four extreme synthetic populations of daily necessities type

Extreme synthetic populations[167]	**Empirical variance**	**Theoretic D&O**	D&O est.	**RG (1)**	**RG (2)**	**RG (3)**	**SRS**	**Two stage prod. and outlet**	**Two stage outlet and prod.**
E. Pure random price ratios that are uniformly distributed, i.e. no product group or outlet effect	3.8	3.6	6.1	6.6	3.7 Good	4.3	3.6 Good	3.6 Good	3.6 Good
G. No random effect in price ratio, only an outlet and a product group effect that are multiplied to total effect	13.1 ±0.6	13.5	13.4 Good	13.7 Good	9.5	7.4	1.7	7.0	7.2
Gy. As G, but 51% of cells lack values	16.6 ±0.8	17.8	17.4 Good	17.9 Good	13.0	11.2	3.6	9.0	9.4
F. Only interaction, see explanation	0.31 ±0.03	0.32	0.90	0.93	0.43 Good	0.63	1.11	0.50	0.58

The theoretical variance according to Dalén & Ohlsson (1995) is in accordance with the empirical variance that was calculated for the thousands of samples drawn from the populations.

It seems as if the variance estimator proposed by Dalén & Ohlsson (1995), D&O, is constructed for a population where there is a clear product effect and a clear outlet effect (population G). For this situation, the estimator is perfect. The estimator also works perfectly when there is lack of observations in the population (population Gy).

In populations in which the price changes occur in a purely random way (E) or in which the measurement process, partly due to the method for replacements and quality assessment, generates data of this kind, the D&O overestimates. If all price ratios have been generated by a random process with independence between outlets and between products, it is best to estimate the variance as if the sample of price notations was a one- or two-stage SRS.

The results for the random groups method RG(1) vary similar to those of D&O est.. These two estimators give the same result for the populations G and Gy, where they are perfect. When D&O systematically overestimates, as in E and F, the RG(2) and RG(3) are better than RG(1).

Table 5 shows the results for populations created from the real KPI data. These populations are made up of many real item strata so that the between-products variance is larger than in reality - if there is any variation between the product strata at all.

[167] 10 outlets sampled from a population of 20 outlets and 12 products (representative items) sampled from a population of 30 products.

Table 5 Average variance estimates (*10^4) for five populations based on KPI data on daily necessities

Populations generated by KPI data	**Empi-rical vari-ance**	**Theo-retic D&O**	**D&O est.**	**RG(1)**	**RG(2)**	**RG(3)**	**SRS**	**Two stage prod. and out-let**	**Two stage out-let and prod.**
GROCERIES4 30 products from cereals, cooking fat, sweets and ice cream for 20 outlets within one retail chain, April 2001. Imputations so that price ratios exist for all 600 cells.	1.42 ±0.07	1.40	1.84	1.94	1.40 Good	1.51 Good	0.72	1.31	0.74
GROCERIES9 As with GROCERIES4, but in September 2001	0.97 ±0.05	0.97	1.29	1.38	0.87 Good	0.88 Good	0.53	0.76	0.69
GROCERIES9y As with GROCERIES9, but for a random sample of 75 % of observations, data have been taken away.	1.78 ±0.09	1.51	2.86	3.31	1.88 Good	2.12	1.58	1.97	2.18
CEREAL4y. 28 cereal products and 24 outlets from one retail chain. Prices are available for 455 of 672 possible cells. April 2001.	0.81 ±0.04	0.85	1.29	1.37	0.85 Good	0.95	0.65	1.21	0.76
CEREAL12y. 29 cereal products and 19 outlets from one retail chain. Prices are available for 335 of 551 possible cells. December 2003.	1.67 ±0.01	1.45	2.51	2.87	1.84 Good	2.02	1.11	1.53 Good	0.91

Table 5 indicates that it is the best to regard the sample as a two-stage sample with a sampling of products in a first stage and independent samples of outlets in a second stage. Here the RG(2) and RG(3) are defined to reflect a two-stage sampling design with products as PSUs and independent samples of outlets for each PSU. RG(2) and RG(3) are closest to the empirical variance obtained from 3 000 samples drawn.

Why do D&O est. and RG(1) overestimate the variance? One conclusion, from table 4, could be that the product and outlet effects are weak and that price changes are generated in a way that could be compared to independent and random for every separate product offer. In 1992, a competition law was adopted in Sweden that forbids nearly all cooperation between enterprises when setting prices. One exception to this is weekly campaigns, where outlet chains may cooperate. Temporary price reductions by 10-30 % for a small number of products during a campaign week is naturally one reason for variations in index estimations. However, coordination of campaigns is not necessarily on a national level. Outlets in the Stockholm region may have one set of products on campaign and other regions may have other products within the same outlet chain. There are also a few multi-outlet enterprises that can have coordinated campaigns across the country.

When we try to explain variation in price ratios in population in CEREAL4y with 23 dummy variables for outlets and 27 dummy variables for products in a regression model, the level of explanation (R^2) is only 15.61 %. Outlet variables on their own have a 6.15 % explanation level and representative items on their own have 9.61 %. When price data are randomly distributed to outlets and products, the level of explanation is between 7.75 % and 14.65 % for 10 such random experiments. In the actual KPI data, coordinated variations between outlets or products are consequently not much larger than for a completely random data.

The estimators D&O and RG(1) give almost similar results. The correlation between results for them are at least 0.97 for the four populations in table 5. The random groups method gives systematically higher estimates, which we have found no explanation for.

Because of the limited set of conditions for this study, there can very well be situations where, for example, one outlet or two in the sample suddenly changes the prices radically, depending on new ownership or the like. As the total sample size is about 50, this can cause a significant outlet effect and an underestimation of variance with the RG(2) and RG(3) estimators.

Clothing and other locally priced items

Clothing and other locally priced items are characterised by the price collector selecting one or more product offers in each outlet for each representative item. In the measurement of clothing, the sample of product offers often consists of 3-5 pieces per outlet and representative item. Furniture is another example, where 2 products offers per outlets are sampled. This means that clothing and furniture are interesting to use as an illustration which will hopefully be applicable to statements about locally priced items in general.

In the following simulations, a sample of outlets and representative items is first drawn. For every combination of outlet and representative item, 2 product offers are randomly selected with SRS. For the D&O variance estimator, the sizes of the population and sample of representative items are considered as doubled, i.e. the two selected product offers are considered as two representative items.

Table 6 Average variance estimates (*10^4) for synthetic populations of clothing and furniture type

Synthetic populations	**Empirical variance**	**D&O est.**	**RG(1)**	**RG(2)**	**RG(3)**	**SRS**	**Two stage prod. and outlet**	**Two stage outlet and prod.**	**Modelbased estimation**
E2A. Population of 20 outlets and 30 items[168]. Price ratio computed as sum of outlet-, item-, interaction-effects and error. Standard deviation for Outlets = Items = Interaction = Error = 1.0	0.076	0.063	0.084 Good	0.063	0.048	0.009	0.038	0.042	0.078 Good
E2A:OIE. As with E2A but without the between-item variance.	0.063	0.068 Good	0.073	0.068 Good	0.070 Good	0.008	0.005	0.063 Good	0.071 Good
E2A:PIE. As with E2A but without the between-outlets variance.	0.042	0.027	0.047 Good	0.025	0.009	0.007	0.039 Good	0.005	0.043 Good
E2A:IE. As with E2A but without the between-outlets and between-items variance.	0.0079	0.0097	0.0128	0.0076 Good	0.0076 Good	0.0050	0.0050	0.0049	0.0081 Good
E2A:E. As with E2A but only with errors	0.0031	0.0042	0.0043	0.0026	0.0030 Good	0.0025	0.0017	0.0017	0.0032 Good
E2C. Populations as E2A with standard deviation similar to furniture data: Outlets=0.60, Items=1.11, Interaction=4.70 and Error=10.32.	0.79	1.06	1.25	0.70	0.73 Good	0.54	0.58	0.50	0.80 Good
E2B. Populations as E2A with standard deviation similar to clothing: Outlets=7.88, Items=4.79, Interaction=11.13 and Error=31.21.	5.9	7.2	8.3	5.4	5.6 Good	3.0	3.3	4.1	6.5 Good
E2By. As E2B, but with gaps for 50 % of the combinations outlet x item.	15.4	19.2	23.0	13.6	14.8 Good	9.1	12.5	13.0	16.3 Good

The model-based variance estimator is clearly the best. It is not completely perfect for populations when there are gaps in the data.

[168] 10 outlets sampled from a population of 20 outlets and 15 representative items sampled from a population of 30 and finally 2 product offers from populations of 10 within each combination of outlet and item.

RG(3) is a random group method approximating a two-stage design variance when outlets are the PSU and products are the within PSU elements. We see, as expected, that the RG(3) estimator fails totally for population E2A:PIE where there is no outlet effect.

Table 7 Average variance estimates (*10^4) for populations based on KPI data for furniture and clothing

Populations created from KPI data	**Empi-rical vari-ance**	**D&O est.**	**RG(1)**	**RG(2)**	**RG(3)**	**SRS**	**Two stage prod. and outlet**	**Two stage out-let and prod.**	**Model based esti-ma-tion**
E2Furniture1. Population created from furniture data for February, May, August and December 2002. 35 outlets and 14 items items[169].	2.8 ±0.2	3.7	4.6	2.7 Good	2.9 Good	1.8	3.0 Good	2.4	2.8 Good
E2Furniture2. A population created from furniture data for 2002, recalculated with the price index. Same size as E2Furniture1.	2.8 ±0.2	3.7	4.7	2.7 Good	2.8 Good	1.8	3.8	2.4	3.0 Good
E2Furniture3. A population with 35 outlets and 10 items created from furniture data for January 2002.	0.9 ±0.1	1.4	1.6	0.9 Good	1.0 Good	0.8	0.7	0.7	0.7*)
E2Furniture3. July 2002	3.3 ±0.2	4.6	5.3	3.2 Good	3.4 Good	2.4	4.9	3.1 Good	4.3
E2Furniture3. December 2002.	4.2 ±0.3	6.3	7.2	4.5 Good	4.8	3.1	5.6	4.2 Good	4.5 Good
E2Clothing1. A population of clothing data from April 2001. At least 4 products offers. 76 outlets and 24 goods items[170]	11.0 ±0.7	15.3	17.4	10.3 Good	11.7 Good	9.1	17.4	14.7	13.5*)
E2Clothing2. A population of clothing data for April 2001, only 39 outlets with a large assortment. At least 3 prod off[171]	16.7 ±1.1	23.7	28.1	16.4 Good	19.4	14.9	26.8	23.9	21.2*)

*) Models could not be estimated for all 2 000 samples

[169] 10 outlets sampled from a population of 35 outlets and 8 representative items sampled from a population of 14 and finally 2 product offers from populations of varying number of products within each combination of outlet and item.
[170] 35 outlets sampled from a population of 70 outlets and 12 representative items sampled from a population of 24 and finally 2 product offers from populations of varying number of products within each combination of outlet and item.
[171] 20 outlets sampled from a population of 39 outlets and 12 representative items sampled from a population of 24 and finally 2 product offers from populations of varying number of products within each combination of outlet and item.

An implication from these results is that it is reasonable to analyse data in the clothing and other locally priced systems as if they come from a two-stage sample with outlets in the first stage (PSU) and products offers in the second stage.

The correlation between the D&O- and RG(1)-estimators for the thousand samples is high, whilst the correlation between these two and the model-based estimator is slightly lower. The model-based variance estimator has a larger variance than the other two estimators, which for example means that the level of coverage for a confidence interval computed as ± 2 standard errors might be lower.

We can note that method RG(3) works satisfactory.

Variance components

The D&O estimator, the estimators based on random groups methods and the model-based variance estimation methods all consist of a number of variance components that can be interesting to analyse. It should particularly be examined how the between-products variance is captured by the between-outlets and by the between-items variance for the D&O and RG(1) variance estimators. Again the D&O and the RG(1) estimators have the same pattern.

Table 8 Average estimated variance components ($*10^4$) for some of the populations in study

	D&O variance estimator			Model-based			
Population	Outlet	Item	Inter-action	Outlet	Item	Inter-action	Product (error)
E2A	0.042	0.019	0.002	0.037	0.033	0.005	0.003
E2A:OIE	0.063	0.003	0.002	0.058	-0.000	0.008	0.005
E2A:PIE	0.005	0.020	0.002	0.000	0.034	0.005	0.003
E2A:IE	0.0049	0.0032	0.0016	-0.0001	0.0000	0.0049	0.0032
E2A:E	0.0017	0.0017	0.0008	-0.0000	-0.0000	0.0000	0.0032
E2C	0.39	0.46	0.21	0.02	0.08	0.18	0.53
E2B	3.5	2.6	1.1	1.9	0.9	0.7	3.1
E2By	8.0	7.6	3.5	2.7	1.8	2.3	9.7
E2Furniture1	2.2	0.8	0.6	0.0	-0.0	1.3	1.5
E2Furniture2	2.2	0.9	0.6	0.0	0.2	1.4	1.3
E2Furniture3 Jan.	0.7	0.4	0.3	-0.1	-0.1	0.1	1.0
E2Furniture3 July	2.5	1.2	0.9	0.7	0.3	0.8	2.5
E2Furniture3 Dec.	3.7	1.5	1.1	0.4	-0.0	1.0	3.1
E2Clothing1	6.6	5.7	3.0	0.7	-0.7	1.0	12.6
E2Clothing2	9.9	9.1	4.8	-1.3	-2.8	4.7	20.9

For sample allocation, it is important to get the right proportions of variance shares for outlets, representative items and product offers. The latter two can often be considered together because a practical and inexpensive way to measure several products per outlet is to define several representative items. Table 8 indicates that D&O overestimates the between outlet variance, which can be explained by the estimation model considering the sample of representative items and products offers as parallel in one stage, and the large variance between product offers is divided to outlets and representative items.

For furniture, the outlet and representative item variances are small whilst interaction gives a variance of the same size as the between-product offer variation. For clothing, nearly all variance is between product offer variances. For sample allocation purposes this would mean that one should collect as many prices as possible in as few outlets and for as few representative items as practically possible.

5 Use of estimated variances

Allocation of resources to the three local price systems

Last year, Statistics Sweden collected data on resources spent at processes such as sampling frame editing, price collectors' travel, price collection, data registration, editing etc.. In combination with estimated variances, an update of the sample design is possible. These are the conditions, where the variances are estimated with method RG(3) on an aggregated level:

Table 9 Weights, costs, estimated variances and best allocation for the three local price systems

	Daily necessities	Clothing	Other locally priced items
KPI-weights	145.9	51.6	245.9
Variable costs[172] (10^3 SEK)	1 102	1 523	3 716
Est. variance for inflation level without fpc[173]	0.037	3.02	0.089
Est. variance for monthly change of inflation with fpc	0.018	2.23	0.041
Optimal allocation of resources for estimating inflation level (10^3 SEK)	667	2 500	3 203
Optimal allocation of resources for estimating monthly change of inflation (10^3 SEK)	620	2 855	2 896

Assuming that we keep the design of each of the three sub-systems unchanged but are able to change the sample sizes proportionally to the resources spent on them, we can find the optimal allocation of resources by minimising a function of three variables under a cost restriction. We now find that the sample sizes should be increased for clothing. We can also see that, if we give priority to change of inflation from one month to another, we would need an even larger sample for clothing than if we set the monthly level of inflation as the first priority. This can be explained by the rapid turnover of products and volatile price changes, which starts in January with the sales.

The results in table 9 cannot be implemented fully because of the large number of retail trade industries covered by the "Other local prices system". We need a minimum sample size for outlets for each industry.

Allocation of resources within the three local price systems

Tables 6 and 7 indicate that the model based variance estimator works well. The analysis variance for clothing and furniture, also the analysis of daily necessities made, clearly show that the variance between outlets and the variance between representative items is small compared to the variance between product offers for an item in an outlet. This and the fact that there is a large cost for price collector's travels to the outlets, leads to a design with few outlets and many products offers, possibly by many items, in each outlet. Therefore, Statistics Sweden have asked all the price collectors what is the maximum number of observations per outlets in all the retail trade industries, considering their own working conditions, the attitudes of the personnel and owners of the outlets. The analysis is made at present.

[172] Costs that are proportional to sample size
[173] Finite population correction

Variance caused by a detailed structure of weighting classes in the KPI

The elementary aggregates in KPI are defined by 287 product groups and 1–6 retail trade industries per group for the goods and services in this study. In all there are 761 combinations of these, each with a fixed CPI-weight. It is a complex task to stratify the populations of outlets and products and to allocate the sample sizes optimally when the total index is a weighted sum of 761 elementary aggregates. With the "half-samples" defined by the RG(3)-method we have computed variance estimates under the model that there are no industry effects, implying that all prices can be aggregated directly to the 287 items. This variance is 15 % smaller for the "Other local prices system", where the number of industries is largest. For daily necessities the "reduction" of variances is only 7 %, here there are only two industries. For clothing, where the variances are very large by nature, Statistics Sweden has, since 1990, applied the "model" that there are no significant differences between industries in price change; data from all industries are processed to produce item indices directly. The samples of product offers is initially made proportional to market shares for the industries selling clothes.

Technical bias in the long term index

Estimates of ratios have a "technical" bias for most sampling designs. For the so called long-term link (1) it is apparent why this bias exist.

Remember that the long-term link for year *y* have product group weights that are values of private consumption during the year *y*, which are price-backdated to the price reference period of the link, i.e. December year *y-1*. This back-dating means that the consumption values are divided by the average of price indicex numbers for the twelve months January to December year *y*. Bear in mind that the index numbers are estimated, some of which with large sampling variances. If a product group happens to get a low price change for December and the whole year on average, the weight gets large by the backdating procedure and the low index gets too large a weight. Index numbers that are large by chance get low weights. The random sampling errors seems to cause a downward systematic error.

This bias was estimated by Statistics Sweden 1996 on purely theoretical basis. Like the Jack-knife-method it is possible to estimate the bias with our RG(3) half-sampling procedure. We have reason to think that the bias is proportionate to variance and consequently inversely proportionate to sample size. If the sample size reduces to half the bias is doubled. This means that the bias of an index based on one of our half-samples is twice the bias of our index based on all data. The bias of our index based on all data can simply be estimated by the difference between the index based on all data and the average of index numbers based on 1 000 half-samples.

Table 10 Estimated bias of the long-term index link for the three local price systems 2001-2003. Percent.

Year	Daily necessities	Clothing	Other locally priced products	All three local price systems
2001	-0.01 ± 0.03	-0.17 ± 0.17	-0.04 ± 0.18	-0.05 ± 0.03
2002	-0.01 ± 0.03	-0.11 ± 0.16	-0.03 ± 0.17	-0.04 ± 0.03
2003	-0.02 ± 0.03	-0.08 ± 0.10	-0.04 ± 0.10	-0.04 ± 0.02

There is a negative bias of the size 0.05 percentage units for the local price systems. These correspond to about half of the KPI. This new finding is a bit lower than the results of the 1996 calculations, which at that time did not lead to any actions by the KPI advisory board.

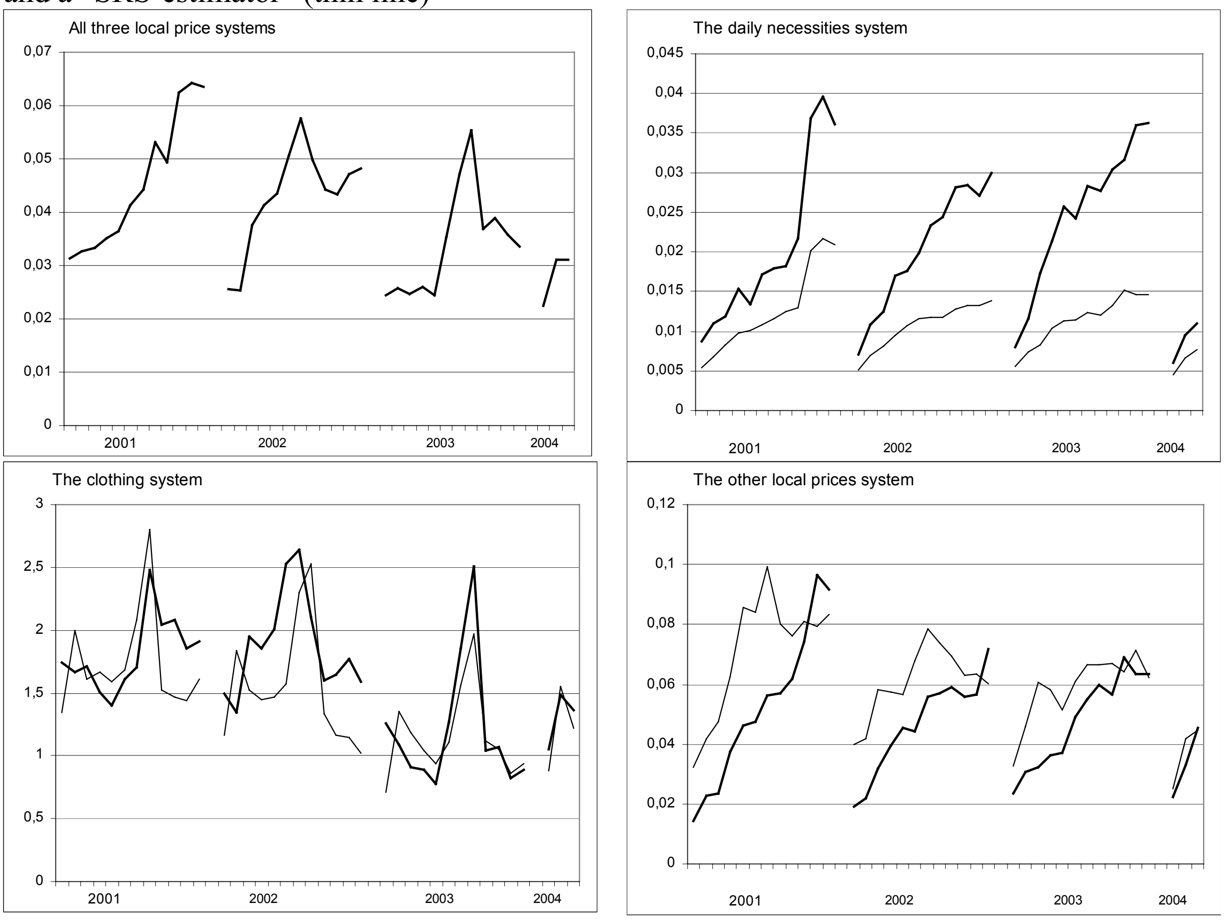
Again, using a smaller number of elementary aggregates, for which the price index can be estimated with better allocation of resources and higher precision, leads to smaller bias – from this point of view. Instead, one can argue that a modelling error is introduced.

Consumer information on variances

Variance for short term index links

There is a significant difference between daily necessities and clothing in the way the variance changes by month during the annual link. The variance for clothing is large already in the first month of the year. This is due to differences in the turn-over rates and frequencies of sales prices. The sample of representative items contain some winter garments, for which we carry forward the April-prices, but there are quite few prices to collect for winter garments in April. This causes a high variance during the summer and a variance from September onwards at the same level as in spring.

Diagram 1 Estimated variances for the short term link ($K_{y-1,12}^{y,m}$) using the RG(3) method (fat line) and a “SRS-estimator” (thin line)

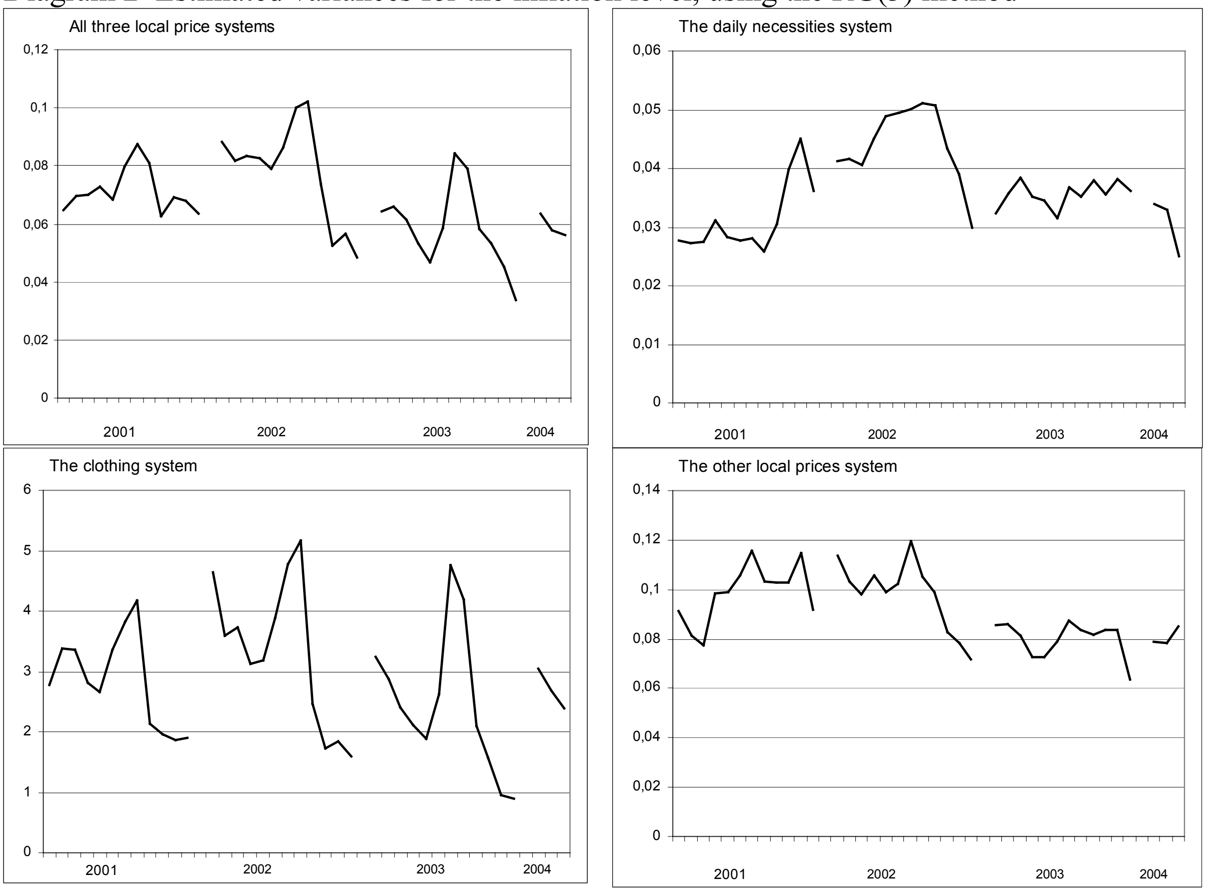


For the three months October to December 2001 the samples of outlets were reduced by about 30 % due to lack of economic resources. This can be clearly seen in the diagrams.

In the diagrams, with a thin line, are shown the "variance estimates" as if we had computed the index by a mean of price changes for all collected data – with no product group weights and no industry weights. This curve is significantly lower than the result for our proposed variance estimator for the daily necessities system. This indicates that there can be some product groups for which the allocation is not optimal.

Variance for inflation level

Diagram 2 Estimated variances for the inflation level, using the RG(3) method

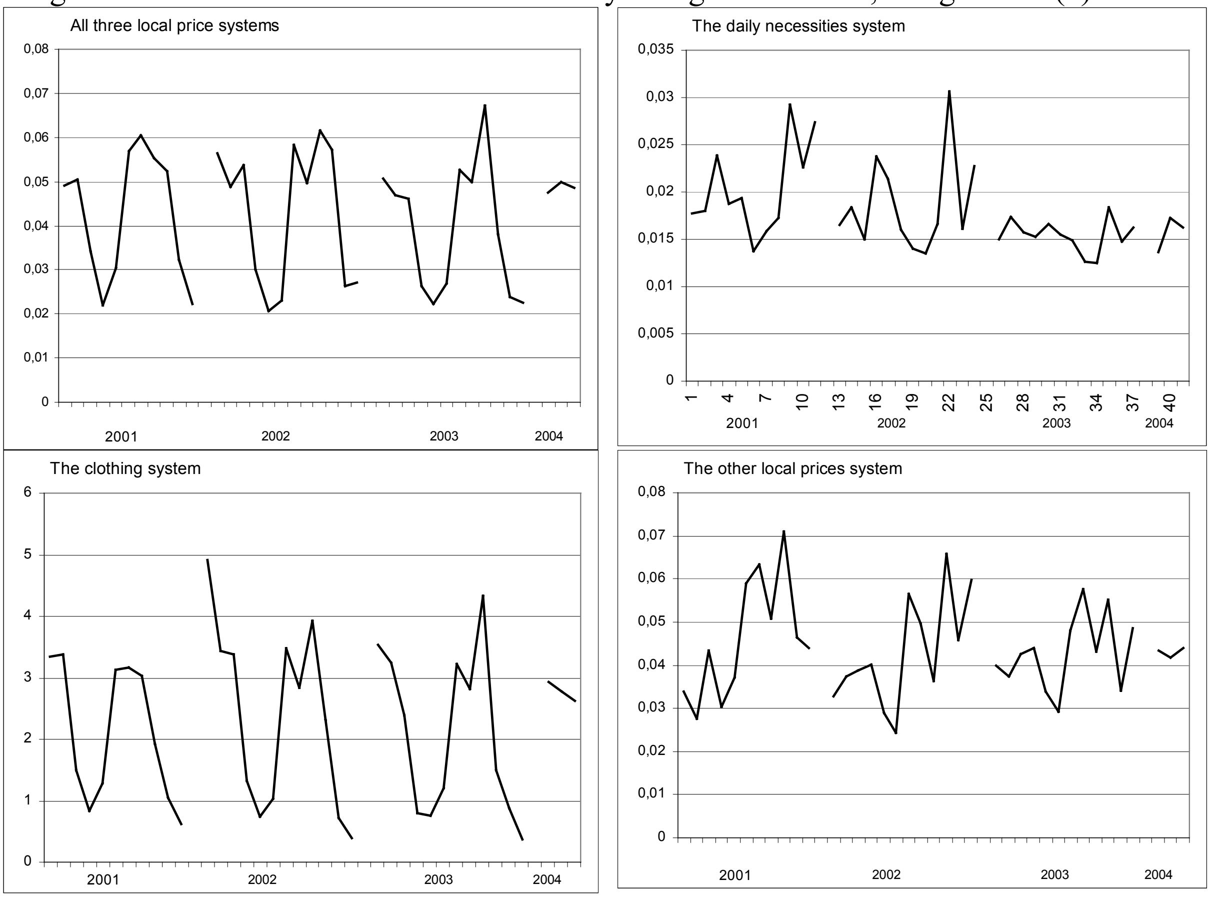


Clothes again deviate from other product groups, there is a seasonal pattern with high variances in June-August. The carry-forward method for winter items is one explanation. There are, probably, some product groups among other goods in the other local prices, for example footwear, where we have the same problem.

The variance of the inflation level is smallest in December. This is due to the fact that this is the only month when the inflation level is equal to a short term index link and for this we have only one sample of outlets and representative items and no update of weights. For other months than December a rough rule of thumb could be to multiply the variance of the December link with 1.5 to get an estimator of the variance of the inflation level.

Variance for monthly change of inflation level

Diagram 3 Estimated variances for the monthly change of inflation, using the RG(3) method



The seasonal pattern for clothing is very clear and it has a significant impact on the total of the retail trade of good and services. The change of the inflation level is approximately the difference between a monthly price change from month *y,m-1* to *y,m* and from *y-1,m-1* to *y-1,m*. If the change of inflation is the most important statistic, more resources should be spent on the clothing survey, as pointed out above.

6 Conclusions and discussion

The use of two-dimensional designs with elements of non-probability sampling methods, annually updated samples and complex estimators creates problems when estimating sampling variances for CPI-statistics.

Dalén and Ohlsson (1995) worked out a variance expression for the Swedish CPI index link in the daily necessities system by Taylor linearisation. They also proposed an estimator for the variance. The daily necessities system has one probability sample of outlets and three probability samples of products, one for each of the three outlet chains that have a dominant market share in trade in daily necessities in Sweden.

A resampling method for variance estimation could be attractive to estimate more complex statistics, such as the twelve-month-change (inflation level) and the monthly change of the twelve-month-change of index. Such methods do not bring more information on the underlying structure of variation, which we need for efficient allocation of resources. For this purpose we have used analysis of variance models. We have developed these models not only to see the structure, but to estimate the sampling variance. This estimator, however, is not practical for complex situations like this.

We have carried out a simulation study to compare three estimators of a random groups type with the Dalén and Ohlsson's estimator and the model-based estimator. We have learned that the following procedure creates reliable results:

Sub-sample a proportion $1/(2-\pi_{(1)})$ of the outlets and a proportion of $\frac{1}{1+\pi_{(1)}\cdot\left(1-\pi_{(2)}\right)}$ of representative items, independently in each outlet, where $\pi_{(1)}$ are sampling probabilities for outlets and $\pi_{(2)}$ are sampling probabilities for products. The latter don't exist for most product groups because of judgemental selection of representative items and must be set rather arbitrary. A large number of such sub-samples, say 1 000, are selected. It is important that these sub-samples are coordinated for all years in study so as to reflect the annual update of samples and sampling probabilities. For each sub-sample the function of index links is computed with the same formulae as is used for the full sample. The mean of squared deviations between the result for the full sample and for each of the sub-samples is the variance estimator.

We have learned that more of the resources should be spent on the clothing survey because the variance is very large. For product groups where the price collector makes a selection of a product offer, the variation in price data can be regarded as a variation between product offers to a large extent and only little variation can be explained by outlet or product group. As the cost of price collector's traveling is substantial, this implies that the samples of outlets should be quite small.

We have seen a significant seasonal pattern in the indices for clothing, especially in the monthly change of the twelve-month-change (inflation). This is partly due to the lack of special summer-garments in the sample and the carrying-forward of a small number of April-prices for winter-garments. Prices for clothing also changes very quickly which makes clothing a bigger problem for precision in the monthly change of the twelve-month-change then in the level of the twelve-month-change. For sample allocation, it is necessary to decide which of these statistics is most important.

7 References

Balk, B. (1989) “On calculating the precision of consumer price indices”, Contributed Papers 47th Session of the ISI, Paris

Balk, B. (1991) “Estimating the precision of a consumer price index; some experiences from the Netherlands”, Contributed Papers 48th Session of the ISI, Cairo. Also presented at the Joint ECE/ILO Meeting on Consumer Price Indices, 18-21 November, Geneva. Modified version in Netherlands Official Statistics 7, 1992, no. 1, pp. 48-49.

Caron, P. and Duval, M-C. (1994) “Estimation of the Consumer Price Index Variance”, Report from Statistics Canada.

Dalén, J. (1992) “Computing Elementary Aggregates in the Swedish Consumer Price Index”, *Journal of Official Statistics*, Vol. 8, No.2, 129-147.

Dalén, J. (1999) “Bedömning av biasrisker i konsumentprisindex (KPI)”, Annex 6 to SOU 1999:124.

Dalén, J. and Ohlsson, E. (1995) “Variance Estimation in the Swedish Consumer Price Index”, *Journal of Business & Economic Statistics*, July 1995. Vol 13, No.3

Norberg, A. (1999): Quality Adjustment – the Case of Clothing. *In Proceedings of the Measurement of Inflation Conference, edited by M. Silver and D. Fenwick. Cardiff University,* pp. 410–426.

Ohlsson, E. (1990) “Coordination of Samples Using Permanent Random Numbers”. *In Cox et al: Business Survey Methods, Wiley*, pp. 153–169.

Ohlsson, E. (1995) “Sequential Poisson Sampling From a Business Register and its Application to the Swedish Consumer Price Index”, *R&D Report 1990:6, Statistics Sweden*

Raj, D. (1968) “Sampling Theory”, *McGraw-Hill*

Rosén, B. (2000) “A User´s guide to Pareto π ps Sampling”, *R&D Report 2000:6, Statistics Sweden*

Shoemaker, O.J. and Johnson, W.H. “Estimation of Variance components for the U.S. Consumer Price Index”, Report from U.S. Bureau of Labor Statistics

Statistics Sweden (2001) “The Swedish Consumer Price Index. A handbook of methods”, http://www.Statistics Sweden.se

Wolter, K. (1985), “Introduction to Variance Estimation”, *Springer-Verlag.*

David Fenwick, Adrian Ball, Jay Beaven: The Quality Adjustment Matrix: a strategic approach to the a priori identification of the most appropriate items in a Consumer Price Index for hedonic quality adjustment.

Abstract:

The dynamics of the market place and the accelerating pace of technological development represent a major challenge to the statistical measurement of the trend in prices of goods and services purchased by consumers. To correctly measure the target universe of transactions associated with the expenditure coverage of the price index, is particularly problematic as it raises a tension between keeping the basket fixed and ensuring that it is up-to-date and representative whilst only reflecting pure price changes. A central measurement issue is the effective treatment of ongoing quality developments of existing products and the proper incorporation of entirely new products. Against this background the paper discusses some of the strategic and practical issues that need to be addressed by statistical offices when reviewing the use of explicit quality adjustment techniques and in particular hedonic regression. Drawing on recent work by the UK Office for National Statistics into explicit quality adjustment methods, the paper considers alternative approaches to identifying at an early stage those parts of the consumer prices index where the application of hedonic quality adjustment can be applied to maximum benefit to the statistical integrity of the index. It explores alternative methods of prioritisation and investigates the possibility of generating a single function that can provide a mechanism for an ongoing strategy for the cost-effective application of hedonics. By providing a reliable set of a priori rules it thus attempts to address the criticism in the Schultze Committee report that the use of hedonics in the past has often lacked strategic focus.

Key words: fixed baskets, quality adjustment, hedonics, Schultze Committee, quality adjustment matrix.

1.0 Background

In 2000 the UK Office for National Statistics launched an extensive research programme to investigate the possible application of hedonic methods for the quality adjustment of goods and services in the Retail Prices Index (RPI) and Consumer Prices Index (CPI). The range of goods studied initially was systematically chosen on the basis of two indicators:

- The relative expenditure share of the good or service in the RPI/CPI basket.
- The associated Implicit Quality Index (IQI.)[174].

The UK approach to the selective introduction of hedonic quality adjustment in consumer price indices may have been more systematic than the targeted intuitive approach[175] employed by the US Bureau of Labor Statistics, but in practice the selection method for hedonic quality adjustment did

[174] IQIs were first developed by Jorgen Dalen and Don Sellwood. They are a measure of the effect of the operational adjustments that have been made to the raw price data in order to obtain the published "quality adjusted" price index. That is the aggregate effect of adjustment, including explicit and implicit methods, to remove non-price effects and arrive at the "true" price change. Relatively large IQIs for specific items may indicate areas where particular attention needs to be given to the quality adjustment techniques particularly where only implicit methods are applied.

[175] Items are selected that are thought a priori to have undergone quality change.

not prove to be completely effective. Only three of the initial fifteen goods that were deemed to be prime candidates for the application of hedonic quality adjustment were selected for implementation in the live index. The goods concerned were PCs, laptops and digital cameras. In addition, only one good- widescreen televisions- was considered borderline. All the remaining eleven goods did not warrant explicit quality adjustment using hedonics[176].

Further investigation of the reasons why hedonic quality adjustment did not prove appropriate shows that whilst the corresponding hedonic functions could be computed, when they were compared with implicit quality adjustment the effect of introduction in the index was not large enough to make implementation cost-effective[177].

Thus whilst it might be true from a statistical viewpoint that "the best candidates for hedonic analysis are categories of goods for which quality change is frequent but incremental and for which the characteristic changes are easy to measure"[178] this isn't necessarily a good criteria on its own on which to base decisions about practical implementation and impact on the index.

It was against this background that the ONS investigated alternative methods of identifying the best candidates for the implementation in the price index of hedonic quality adjustment. The aim was to try to differentiate between those goods that have sufficiently high levels of quality change to warrant explicit quality adjustment and those that don't , where implicit quality adjustment is likely to be problematic. This can then be used as an initial filter prior to the consideration of hedonic quality adjustment.

2.0 The investigation

Investigations focused initially on two primary quantitative measures relating to quality change:

- The turnover in items as represented by the number of price quotes that disappear from the basket.
- Implicit Quality Indices (IQIs).

However, IQIs can be time consuming to compute and have not proven to be totally reliable indicators in the past of the need for explicit quality adjustment and are, in any case, of limited use as they can only be calculated for items already in the index. In addition, a further study of the retail market suggested that turnover rate as a concept is too simplistic and on its own is a very blunt tool. These considerations led us to consider an alternative two-dimensional matrix approach that is characterised by two statistics:

- Rate of turnover of models – are significant proportions of older models disappearing from the market and being replaced by new ones?
- Rate of technology change in the good- are new features and improvements regularly being introduced?

[176] Fridge-freezers; dishwashers; VCRs; radio-cassette recorders; DVD players; still cameras; vacuum cleaners; audio systems; microwave cookers; mobile phone handsets; Widescreen TVs.

[177] In the cases of VCRs and 14" televisions relative sales declined as a result of being displaced from the market by DVD players and by small televisions, combined with a VCR or DVD.

[178] A direct quote from the Schultze Panel on Conceptual, Measurement and Other Statistical Issues in Developing Cost-of-Lining Indexes ("At What Price? Conceptualizing and Measuring Cost-of-Living and Price Indexes"). By way of illustration the Panel offered personal computers which are seen as an obvious candidate whilst cars [where, for example, measuring performance is highly subjective] and clothes [the challenge of fashion] are seen as problematical.

Clearly if an index item is subject to both a high rate of turnover and of technological change then it is likely to be a prime candidate for explicit quality adjustment. Similarly, the index compiler would not want to give priority to explicit quality adjustment where an item has both low turnover and technological change. Both statistics in the matrix potentially can be calculated from the raw data used to compile the CPI or from scanner data but the CPI data is only available for goods already in the index and its value is constrained by the amount of detailed characteristic information which is collected (usually a lot less than for scanner data). It was for this reason that the study was based in most part on an analysis of scanner data[179].

	Low rate of technology change	High rate of technology change
Low turnover rate	NO	?
High turnover rate	?	YES

3.0 The results

Turnover rates

Two measures of turnover were investigated:

- *the rate of models leaving the market.* This is defined as the proportion of models that have disappeared from the market in the following month. It is closest in concept to observed turnover rates in the RPI sample itself.
- *the level of churn in the market* . This is constructed by pairing successive months, and determining how many models are only available in one of the months and expressing this as a proportion of the total number of distinct models available in both months

The rate of models leaving the market allows us to directly compare measures derived from RPI and scanner data respectively. It is also of more direct relevance to the reliability of our current implicit method of quality adjustment, since the latter is a function of the number of models that continue to be available in the shops from one month to the next. The churn rate is of limited relevance in this respect, as the arrival of new models has no effect on the reliability of implicit quality adjustment. However, the churn rate gives a better indication of the underlying dynamics of the market place and can therefore be said to better represent overall the position relating to technological change. Chart 1 shows the monthly turnover rates of models leaving the market for the goods studied. When interpreting this chart it should be noted that, firstly, the reference year varies between goods reflecting the data to hand and secondly, that the volatility of the data plot for PC's results from the use of the relatively small RPI sample which is much reduced compared with scanner data.

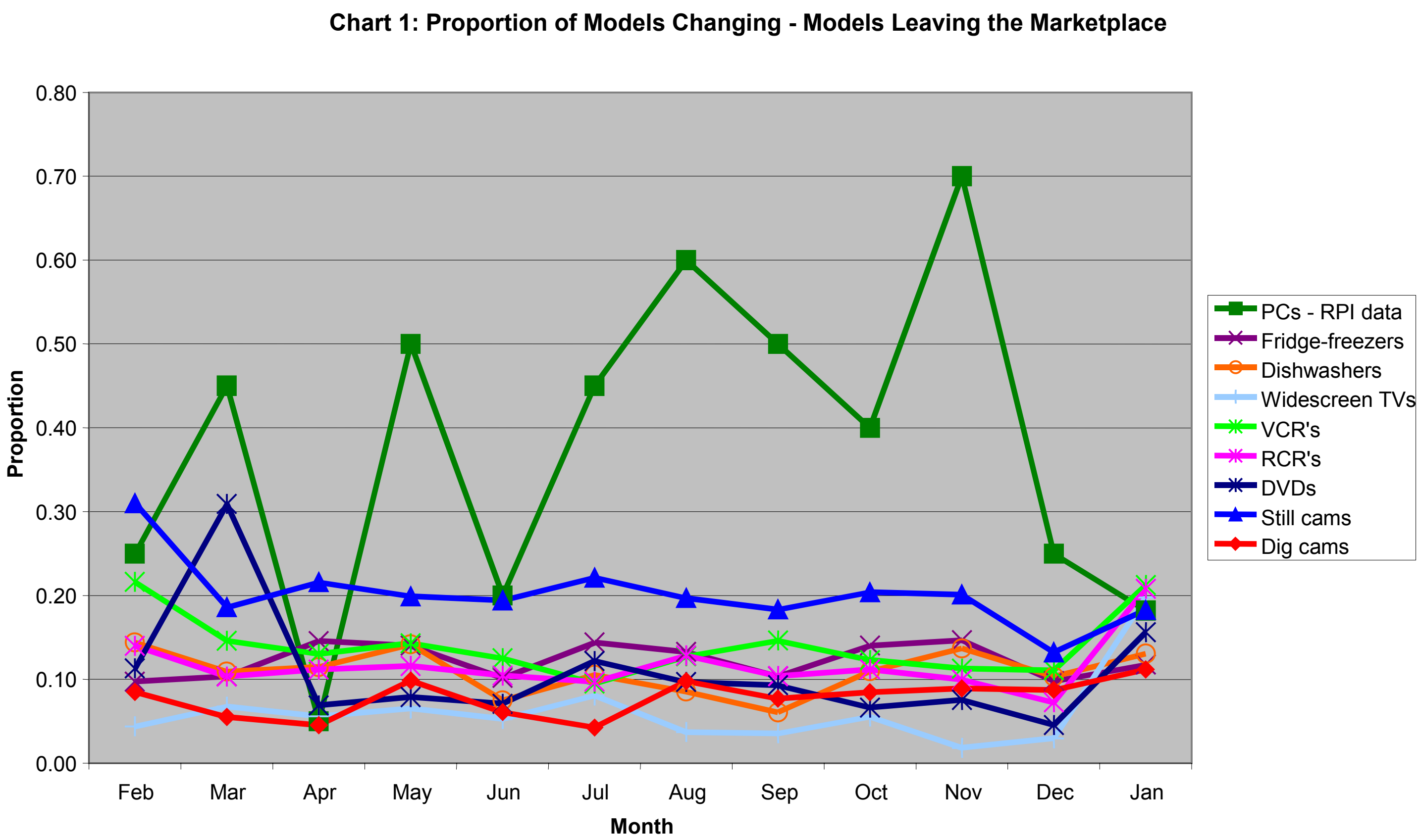
Three points emerge:

- The special position of PCs. The turnover rate is about twice as high for PCs than for any other good studied.
- Turnover rate appears to be a relatively blunt instrument when used alone in determining the need for explicit quality adjustment. Thus it can be seen that the variation in turnover rate is

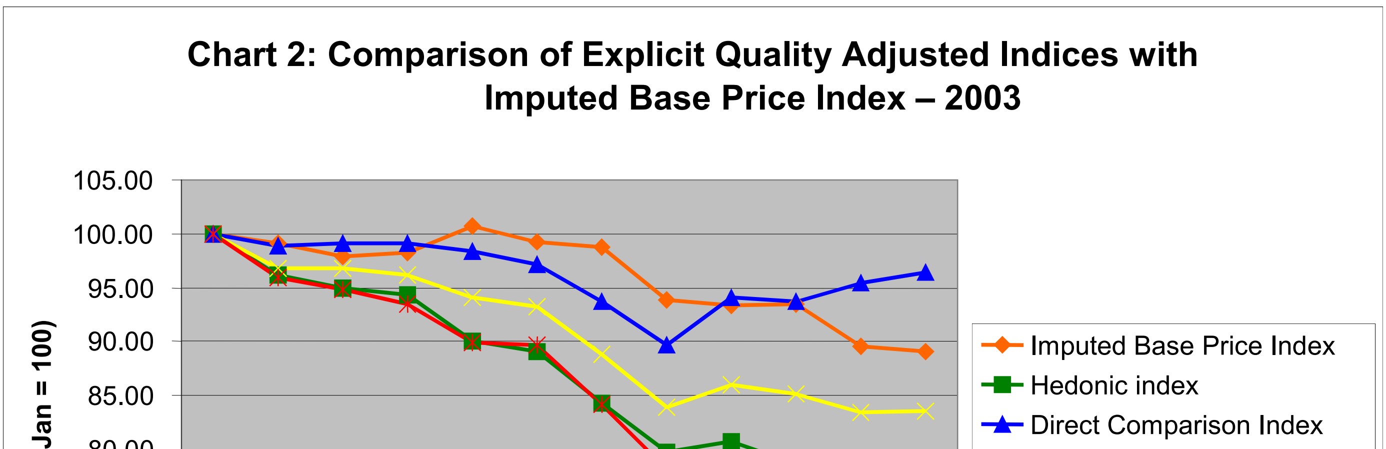
[179] The exception was PCs where models, their characteristics and prices, are collected from the Internet due to the lack of detail in the available scanner data.

relatively small between the goods studied although the turnover rates are not small in absolute terms.

- The turnover rate is not on its own a reliable indicator of the need for explicit quality adjustment. For example, although the turnover rate for digital cameras is the second lowest at seven per cent, further investigation lead us to introduce hedonic quality adjustment for this good.



Further analysis indicates how a high turnover rate can be strongly associated with poor performance of implicit quality adjustment methods. This is illustrated in Chart 2 which compares the hedonic quality adjusted PC index with both the corresponding direct comparison index and the corresponding imputed base price index..

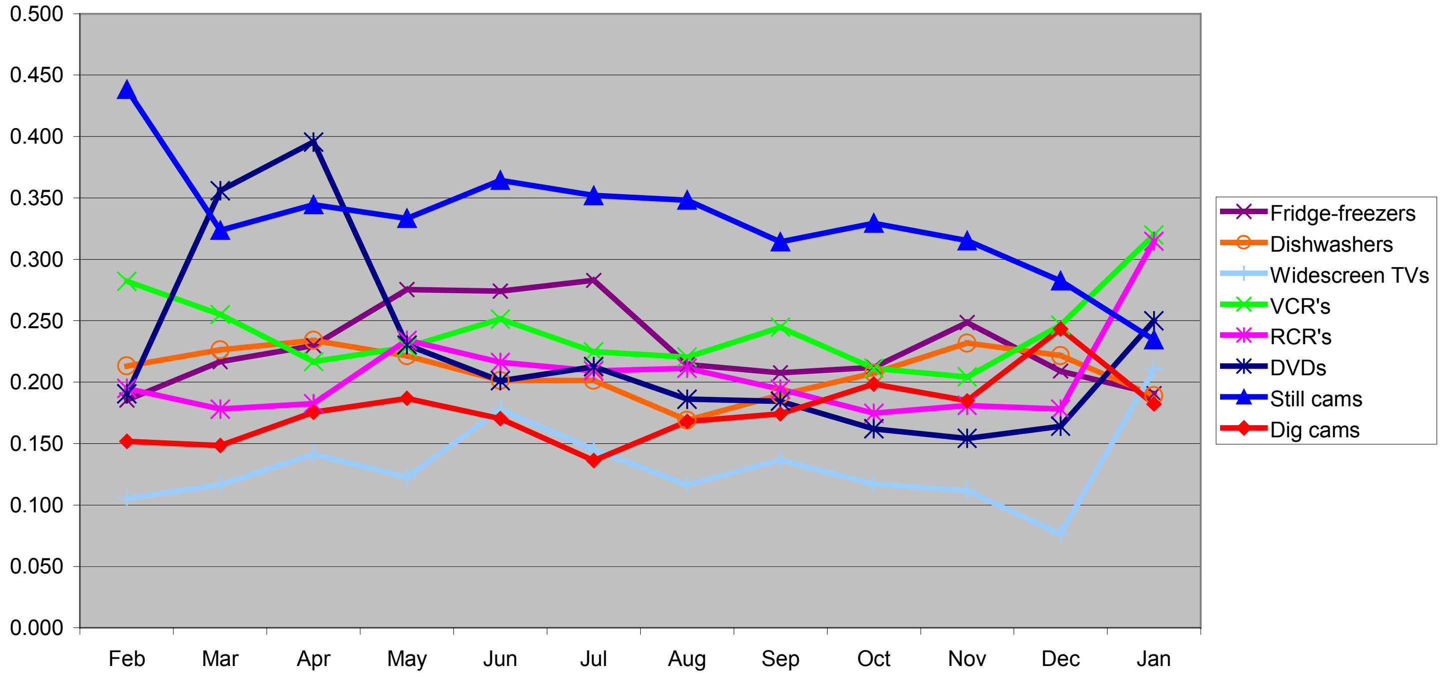


PCs apart, the corresponding levels of churn in the market is shown in Chart 3[180]. As might be expected in a broad state of market equilibrium, the churn rate is about twice as high as the turnover rate.

It is instructive to note that, with the exception of still cameras, the picture to emerge is very similar to that which emerged from an analysis of turnover rates.

[180] PCs have been omitted due to volatility in this measure resulting from the relatively small sample.

Chart 3: Proportion of models changing - Scanner Data - Models entering and leaving the market



Rate of technology change

The measurement of the rate of change in technology was more problematic. Early work focussed on examining the rate of emergence of new features for models but this was soon abandoned as being only partially informative. The next thread of work examined the relationship between levels of attributes and the year of introduction of the model. This produced interesting results for digital cameras, indicating that measures of attribute change can yield suitable results. In particular looking at the attribute megapixels over time showed how the technology had changed significantly in the past seven years. However, it could not generally be applied to other items because of a lack of information on year of introduction of a model. Because of this, efforts were re-directed at developing an "index of attribute change" based on a chained measure of the change in the level of attributes between goods leaving the market and those entering the market, with the attributes that significantly influence the price of the item being identified using hedonic regression methods applied to scanner data[181].

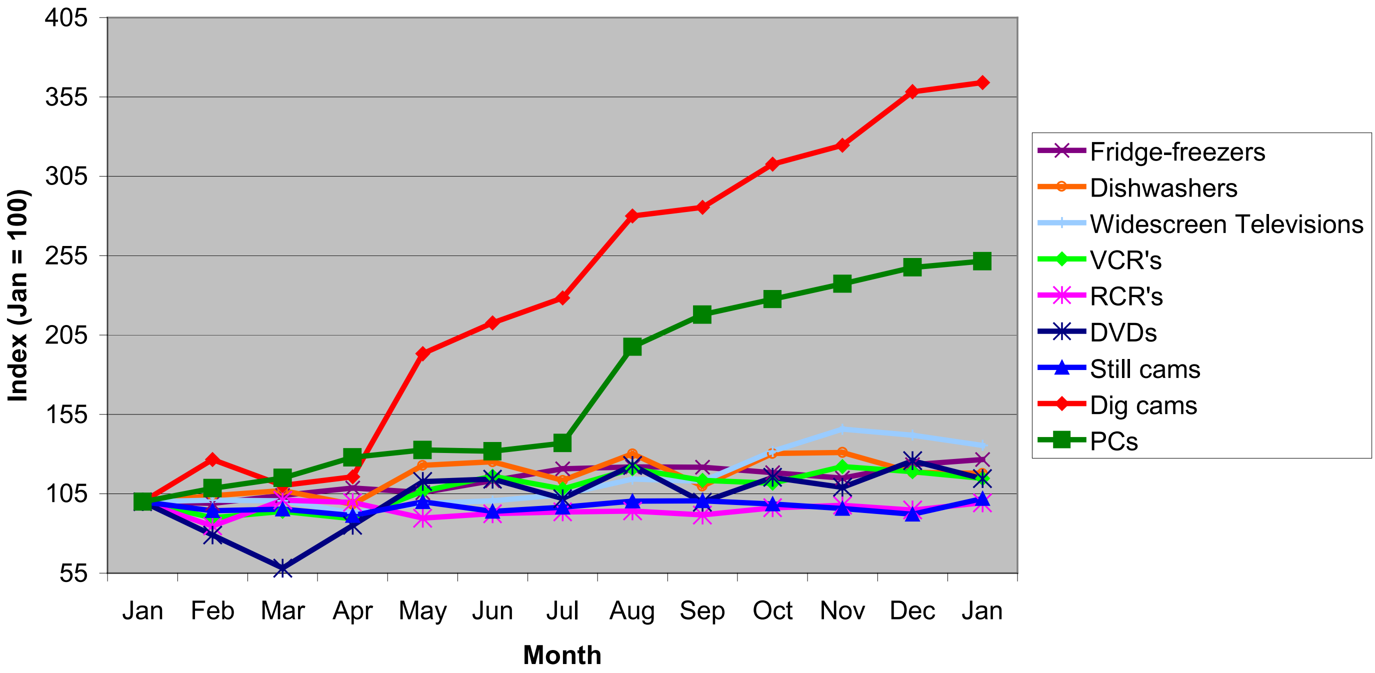
From the same scanner data, the average turnover (entering or leaving the market) of the attributes so identified is then calculated. For numeric variables the latter will be a straight average whilst for discrete features the calculation relates to the proportion of models containing the feature. An

[181] For instance, hedonic regression shows that for televisions the important price determining characteristics are a mixture of discrete and numeric variables: Screen Size; Flat Screen; NTSC tuning; Dolby sound; DVD player in built; PC in built; Real flat screen; Digital tuner;100 Hertz picture.

overall weighted index is then calculated by taking a weighted average of the log of the turnover in price sensitive features (as calculated above), using as weights the t-values for the attributes from the regressions[182] and expressing this in index form with January =100. The index was computed taking into account whether the attribute represented an increase or decrease in quality - the "net rate of attribute change"- and also in absolute terms where a change in features is treated in the same way regardless of the direction of the quality change - the "gross rate of attribute change".

It can clearly be seen from Chart 4 that the "net rate of attribute change" is a strong driving factor in determining the potential need for explicit quality adjustment. In particular, there is a strong demarcation between the two goods identified in our earlier research programme as being suitable candidates for hedonic quality adjustment (and where the method was subsequently introduced) and other items which were researched but where hedonic quality adjustment was not found helpful. The two goods concerned had experienced rates of attribute change two to three times higher than for other goods.

Chart 4: Attribute change Index - all goods



We can conclude that, not surprisingly, it is the level of change in price determining attributes that dictate the need for explicit quality adjustment rather than the turnover in models and that a high turnover of models is not necessarily associated with a high turnover in attributes.
This is backed up by our previous work which led us to introduce hedonic quality adjustment for digital cameras despite their relative low rate of "model" replacement. The question then is whether a threshold can be set a priori to differentiate between those items where hedonic quality adjustment should be applied and those where there is no strong justification.

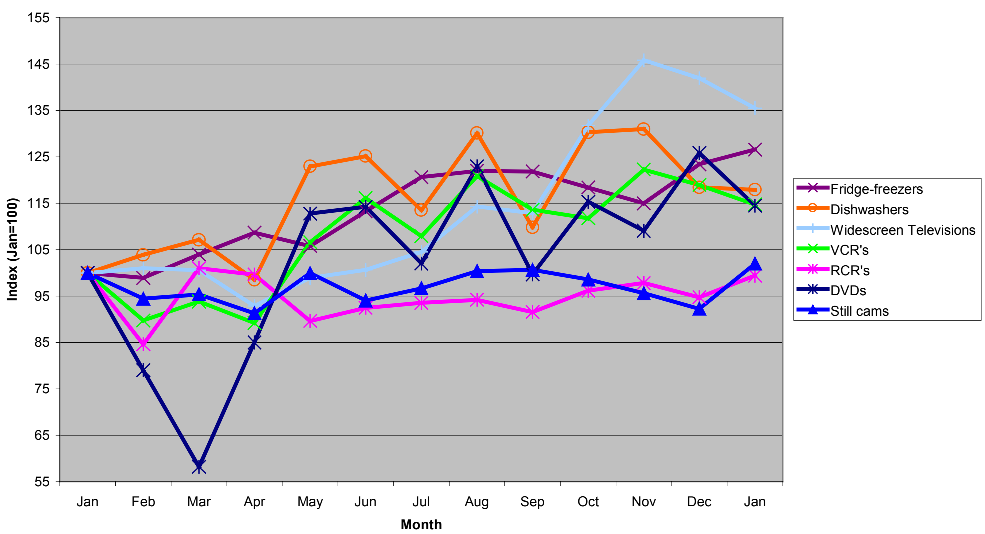
[182] The t-values indicate both the importance of the variable and whether a change in the attribute increases or decreases the quality of the item. For example the attribute 'volume' for digital cameras has a negative t –value, so the heavier a camera the more quality decreases.

A further look at the results- Chart 5 redraws to a larger scale the data in Chart 4 but excludes PCs and digital cameras- leads to the following observations:

- There is very little change in price determining characteristics for radio-cassette recorders and still cameras (film cameras). For radio cassettes, this re-enforces an earlier research conclusion that hedonic quality adjustment added little value to index compilation as there was very little real change in the market place. For still cameras the situation is more complex as earlier research indicated that the market continued to be innovative but that hedonics was unable to capture the quality change. This latest work seems to indicate that the problem may be that the available scanner data excludes important price determining characteristics or that hedonics has not been successful in identifying them.
- All the other items studied, wide screen TVs apart, showed an increase in the rate of change in price-determining attributes, although there was a blip in March for DVDs and there was some volatility associated with dishwashers[183]. Earlier research had indicated that the use of explicit (hedonic) quality adjustment was not justified for these goods even where there was a high turnover rate of models, such as with VCRs. As mentioned, dishwashers and DVDs demonstrate a volatile pattern in attribute change over the year
- There is the same slow increase in attribute changes for widescreen TVs as for most other goods but with major changes occurring in October and November. Earlier research into the application of hedonics to this good suggested that hedonic quality adjustment added very little value to the statistical integrity of the index because of the relatively low turnover rate in models. But clearly such conclusion may need to be re-visited if the turnover rate in models is volatile and becomes significant in particular months. The issue to be resolved is the level at which turnover becomes a significant factor.

A corresponding analysis of the "gross rate of attribute change" is shown in Chart 6 and suggests

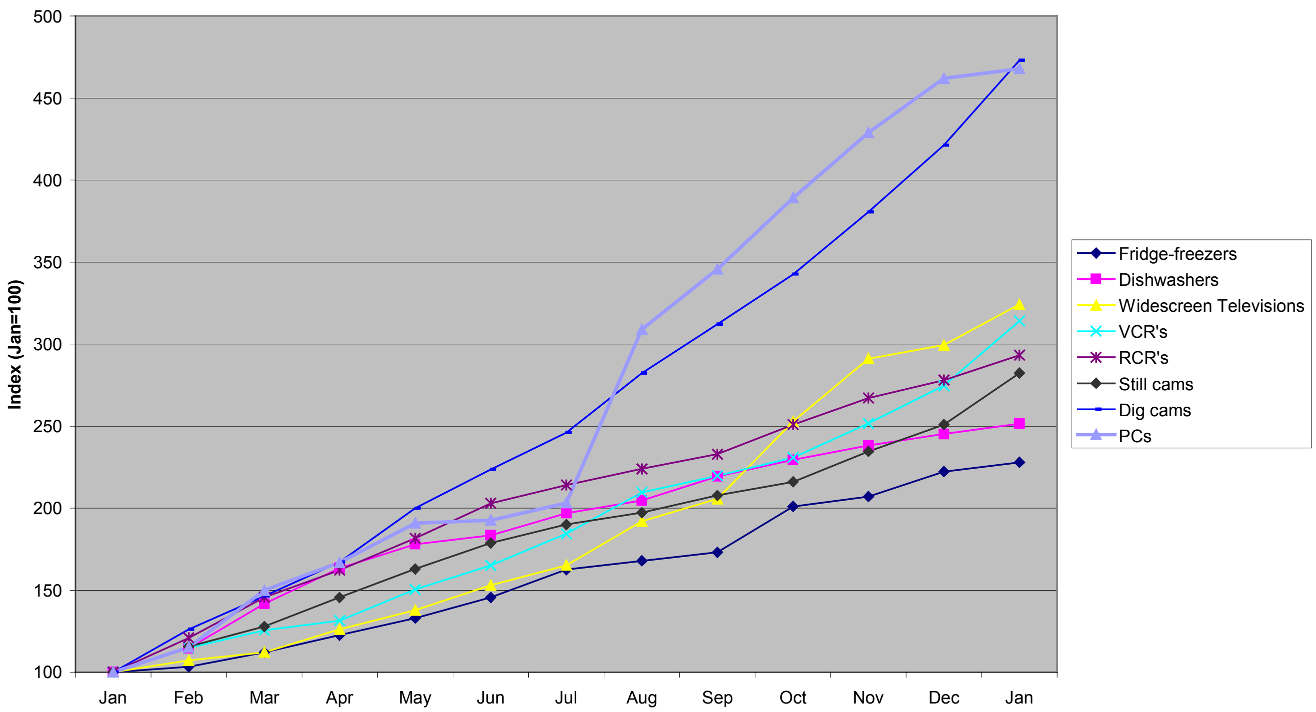
Chart 5 Attribute Change Index - Excl PCs and Digital Cameras



that the gross measure is not particularly helpful in terms of its ability to differentiate between goods that would benefit from explicit (hedonic) quality adjustment and those that wouldn't. This is because it disguises the fact that a lot of quality change might occur but the increases and decreases in quality might cancel one another out.

[183] In both cases this is due to limited number of models available.

Chart 6: Absolute quality change



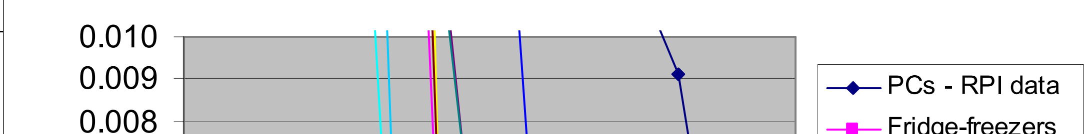
Combined measures

The final step in this research was to explore the potential for compiling a single indicative measure of the a priori need for explicit (hedonic) quality adjustment based on a combination of the measures discussed above. Cluster analysis was also considered but dropped because of the relatively small number of observations involved.

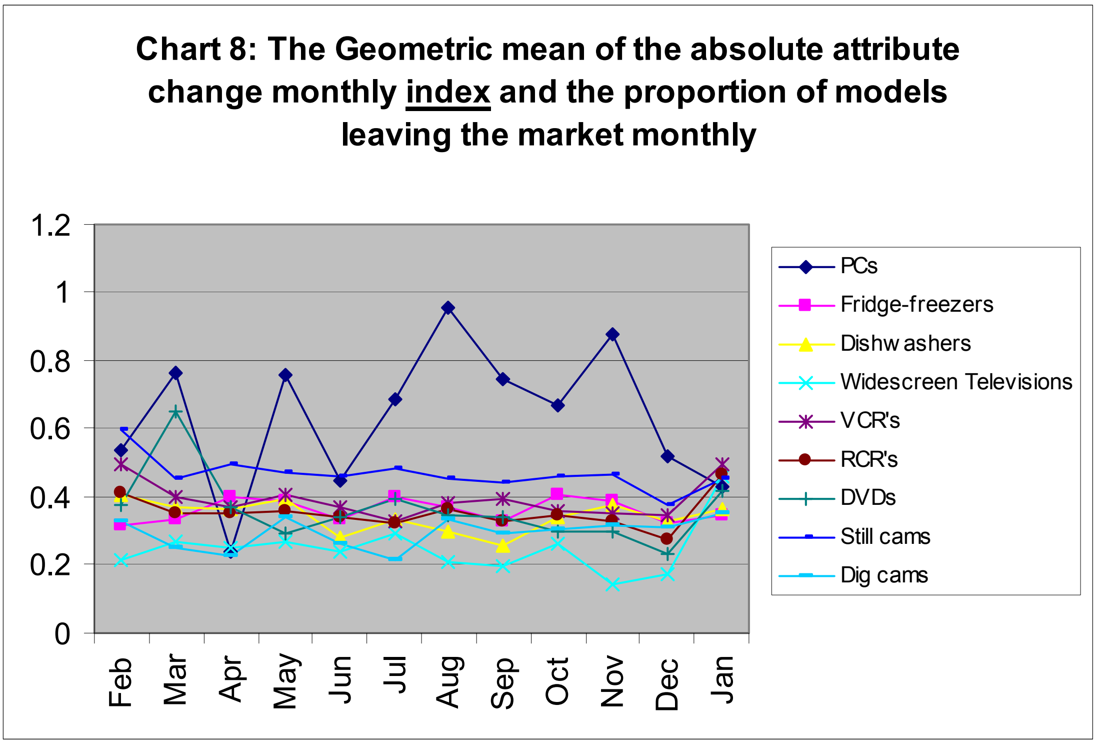
A variety of combined measures were computed.

Chart 7 shows the results of a measure based on multiplying the absolute attribute change monthly index[184] with the proportion of models leaving the market monthly. Because the resulting measure is small in numerical terms the graph focuses on the extreme lower end of the index. The further to the right an item appears on the graph, theoretically the greater the likely need for explicit (hedonic) quality adjustment because the combined measure is higher and falling at a slower rate. Putting to one side the conceptual basis of such a calculation, it can be seen at a glance that this combined measure does not increase the ability to pinpoint items that would benefit from explicit quality adjustment- indeed in some respects this combination is less successful. This is illustrated by digital cameras, which are hedonically quality adjusted in the RPI and CPI and which clearly require quality adjustment when looking at a technology change measure, but has one of the lowest combined score on the chart. Unsurprisingly, plotting the geometric mean of this combined measure as shown in Chart 8 to dampen the extreme values doesn't assist the visual presentation.

Chart 7: The absolute attribute change monthly index multiplied by the proportion of models leaving the market monthly



184 … in the base
m…



Finally it should be noted that there was no significant advantage in re-computing this combined measure using the alternative concept of total change in the sample (new models plus old models) instead of the proportion of models leaving the market.

4.0 Concluding remarks

Research to date suggests that an analysis of the turnover rates in price-determining characteristics has the ability to differentiate between goods that are clear candidates for explicit (hedonic) quality adjustment and those that are not but that there is no easy method of identifying more marginal cases or cases where the rate of model change might be at a level where explicit quality adjustment is desirable even though the rate of change in attributes indicates that the position is marginal. The combined measures that have been tested do not increase the power to discriminate although an alternative option to pursue for future research might be some form of weighting optimised to maximise the predictive power. Finally, it is noted that the compilation of a *net or gross rate of attribute change* involves scanner data and hedonic regression (for the identification of attributes) so operationally does not represent a significant departure from the traditional “test and see” approach of running full pilot studies for individual goods.

Yoel Finkel and Dorit Zioni[185]:Automization in the CPI, quality assurance and EFQM

1. Abstract

The Israeli Central Bureau of Statistics (ICBS) has redesigned, in recent years, its procedures for computing the CPI. These included introduction of new technology and automization of fieldwork and editing processes. An important factor in the new system is process control and auditing. In order to estimate the "intuitive" improvement in the CPI, the Consumer Prices Division is in the midst of assimilating EFQM methods, including quality indices of the CPI procedures. The purpose of this paper is to present the methodology that was developed to enhance the validity of the CPI procedures and to discuss the relationship between the EFQM model and the CPI. Quality indices may be constructed through this system in order to track future improvements in the CPI. Portions of this paper will be based on Chapter 12 *Organization and Management* of the new international manuals on CPI and PPI. This coincides with an additional objective of the Prices Department at ICBS: to align the Israeli indices with many of the procedures presented in these manuals[186].

2. *Introduction*

In 1996 ICBS initiated the CPI Development project, which was mostly completed at the end of 2001. During the project, ICBS conducted a joint examination of the three intertwined surroundings of the CPI: methodology, technology and field operations- and their connection to the structure and organization of the professional units responsible for construction, collection, computation, analysis, publication and dissemination of the CPI. ICBS believed that the first step in solving many of the complicated issues of the CPI was to build anew the infrastructure of the index, using advanced technological means. The technological and organizational features of the new system were designed to enhance the methodology of the index and the preceding field operations. Therefore, the system has two major and connected parts: the price collection sub-system and the price compilation and analysis sub-system.

---

[185] Central Bureau of Statistics, Israel. The authors thank Ms. Svetlana Gluzman and Ms. Merav Yiftach from the Israeli Central Bureau of Statistics for their assistance in preparing this paper. Any mistakes are the sole responsibility of the authors.

[186] Another paper in the series of issues examined within this framework was first presented at the ECE-ILO meeting on Consumer Price Indices, December 2003, Geneva. "The Treatment of Seasonal Products in a CPI" reviewed the

The price collection sub-system structure is based on a wide area network communication system for ICBS regional offices and operations using handheld and personal computers. Connected to this network are the price collectors (field and office), field supervisors and the commodity analysts. The tasks performed in the collection sub system include detailed scheduling of price collection to each price collector, price collection and automatic logical checking at the micro and macro level. The principle behind the automatic logical checking at the collection stage was to move editing, logical and statistical checks, from the headquarters to the field operations- as close as possible to the reporting occurrence.

The prices compilation and analysis sub-system structure is based on a local area network communication system for the CPI headquarters. Connected to this network are the officials providing subject (commodity analysts), methodology, publication and information services. The tasks performed in this sub-system include: planning, quality control, final nesting of price observations (already filtered in the collection process), index computation, analysis, research and publication.

The project has brought about new modes of cooperation and improved working procedures. Analysis, planning, development and assimilation of new technologies have led to already reaped benefits of methodological analysis.

In this paper we will examine the automatic logical checking in both sub systems- collection and prices, and attempt to build quality indices to measure the methodological improvement already achieved or to be achieved in the future. In addition, we will review other quality checks made by officials from both field operation and CPI analysts. Section 3 describes the quality control process enabled by the automization of the CPI; section 4 presents quality indices devised in the measurement process of the EFQM framework and in section 5 a brief summary of the paper and possible work for the future are presented.

3. Quality checks in the CPI

These include checks at various stages in the index compilation process: fieldwork, editing and computation stages.

---

feasibility of the methodology introduced in chapter 22 of the manuals, in a "real" CPI, by simulating actual price and quantity data from the Israeli CPI in the seasonal formulae.

3.1 Quality checks in fieldwork

Checking processes begin in fieldwork. The fieldwork is the first stage of current CPI compilation and is a major infrastructure of the index. An important principle to achieve in the automization process is to have editing, logical and statistical checks as near as possible to the reporting occurrence in order to minimize errors at later stages of the index computation. Here we have two kinds of checks at fieldwork- logical editing of data during price collection in the software designed for the hand held computers and back checking of fieldwork by the survey division officials in the main or regional office[187].

3.1.1 Automatic checks

3.1.1.1. Price changes - The price entered is compared with the price for the same defined chosen product in the same outlet in the previous month and the price collector is queried when the price change is outside preset percentage limits. The limits may vary according to specific products or consumption groups. The percentage limits were determined by looking at historical evidence of price variation[188]. If there is no valid price for the previous month, for example, because the produced good was out of stock and no transaction could be made; the check is made against the last "real" price entered for the same chosen product.

3.1.1.2. Maximum/minimum prices- A query is raised if the price entered exceeds a maximum or is below a minimum price for the group of goods or services of which the particular product is representing. The range is updated every three months and is derived from the validated maximum and minimum values observed for that price transaction in the previous month expanded by a standard scaling factor. This factor varies between price transactions.

3.1.1.3. Comprehensiveness checks

Data completeness is vital for the calculation of the CPI. For each chosen product there may not always be many data required, but all data are vital for calculation and identification of the chosen product at the following month. While the most important data for the index are the prices, if the price observation is lacking some additional data, identification of the product may be insufficient and would require some form of imputation. Therefore, comprehensiveness checks are intertwined inside the system to avoid skipping fields with vital information, like price, product descriptions, etc.

[187] For a detailed discussion on back checking refer to the Manual on CPI, pp. 219 or the Manual on PPI, pp. 311-312.

3.1.1.4. Chosen product description

Accurate price transaction descriptions are critical in ensuring price transaction continuity. Descriptions should be comprehensive to ensure that the price collectors or reporters can price the same chosen product in each collection period. They must record all information that uniquely defines the chosen product selected. Accurate chosen product descriptions assist the price collector, respondent, and CPI staff in choosing a replacement for a chosen product that has been terminated and also helps to identify changes in quality. The technological system assures better accuracy and transparency of the chosen product. It enables the price collector to bring much more and accurate information, regarding the sampled product, than he or she could in paper questionnaires. While in the paper questionnaire, there was a limit to the number of characteristics and attributes, in the computer all these are in the database and are shown to the price collector. In addition, each change made in the chosen product, is automatically marked by an "item remark" as a quality change, non quality change, correction/completion, temporarily/ seasonally ran out, permanently ran out, or unavailable. If there hasn't been any change in the chosen product, it will be automatically marked as "no change".

3.1.1.5. Logical editing for running out of items

After three months that the chosen product was in status of "temporarily ran out", a query is raised to remind the price collector to make sure that the sampled product is back in business. If it isn't, he should either replace the sampled product or change its status to "permanently ran out". The options in the item remarks will be "no change", "quality/non-quality change" and "permanently ran out". After three months that the product has been in the status of "permanently ran out", a query is raised to remind the price collector that he should either choose a new product for the defined item, or mark it as "unavailable". The options of item remarks in that case will be "new product" or "unavailable".

3.1.2 Back checking and process auditing

The second type of checks is a beck check at the end of each collection period and process auditing. The checks are not automated[189] but performed by the survey division's officials, who function as the auditor of price collection, as another important tool for improving and monitoring data collection. There are four kinds of auditing and back checks in the Israeli CPI:

---

[188] The percentage limits for seasonal products like fresh fruits and vegetables are 30%, for seasonal clothing and footwear 20%, for non seasonal products 15% are for regular services 10%.

[189] However part of the process is based on the information provided by the database of the automatic process of price collection.

1. The auditor occasionally accompanies collectors during fieldwork- whether data are collected by phone (CATI) or personal visits (CAPI). The auditor actually ensures that the price collector is following proper procedures and instructions and performing the collection competently. The auditor also functions as a trainer, which means that he (or she) gives some coaching to correct any errors. After returning to the office, the auditor is required to file a standard report detailing his observation. This report includes findings, problems encountered in the collection process, and a recommended course of action. Auditors may advise that a collector receive extra training on certain aspects of price collection, raise general problems where solutions need to be disseminated to all price collectors, or revise existing instructions or procedures.
2. The auditor occasionally replaces the permanent collector in the field. He actually collects prices from the work portion of each price collector once a year. During price collection the auditor assess the competence of the price collector, identifies common mistakes and evaluates the need for extra training, validates that he follows procedures and instructions of price collection, identifies areas where collection is problematical, and most importantly, reviews the comprehensiveness of the description of the chosen product. At the end of price collection, the auditor compiles a report detailing his observation as mentioned in paragraph 1 above.
3. Another way of monitoring the standard of price collection is to carry out a back check – a retrospective check of fieldwork at the outlet level. At the end of each collection period the official from the survey division addresses 10% of the outlets in the sample of each work portion of each price collector, to make sure that the price collector actually visited the outlet and in addition to see to that the chosen products which were marked as “unsold” or “temporarily unavailable” are actually unavailable or not sold at the outlet.
4. In addition, at the end of each collection period, the official from the survey division calls all the outlets, in which price collection was unattainable for some reason (the outlet was closed, the owner refused to hand out the prices etc.), to confirm the information and to see if there is a possibility to collect prices from them.
5. Another mean for auditing is also a back check of the solutions given to the computer queries by the price collectors. The official from the survey division follows the reference of the price collector to the automatic logical editing of data. The tool he uses to do so is a special module built for follow up and control. For each price collector the auditor perceives the type, the number, and the percentage of queries rose during price collection. At the end of each month the auditor compiles a report, which sums up the treatment of each price collector. Feedback is given to price collectors according to the results of the quantitative

data. Thus price collectors recognize the value of each logical editing of data, each logical editing of data gets the appropriate attention, and the quality of the data is improved.

3.2 Quality checks by CPI head office staff

There are several types of editing and statistical process controlling carried out by the commodity analysts from the CPI head office staff. The two main categories to be discussed below are (i) weekly checks on the prices and description that are fed into the database on a daily basis by the CAPI and CATI price collectors and (ii) monthly checks on the indices of items, consumption groups and aggregated CPI[190].

3.2.1 Weekly editing by CPI head office staff

Since the computer system enables more efficient editing than a paper based system, one dozen quality checks are performed during the weekly editing process. In order to ensure consistency in the editing process, all analysts are trained to use the queries and to provide identical solutions for same situations. The training process is a continuous one with the relevant officials meeting at a set time each week to discuss the more complicated issues encountered during the editing process. The computer supported quality checks in the editing process are:

1. Double check on all automatic substitutions of items
2. Check on items whose attributes have not changed and even so were determined to be non-comparable
3. Check on new items that were deemed comparable to the former ones
4. Check on "non-quality adjusted" substitutions that were deemed comparable.
5. Check on substitutions that include remarks on sale or other special price.
6. Check on substantial price changes that arrived without a remark from the price collector
7. Check on vat. and exchange rates
8. Check on price level compared to other observations
9. Check on price scaling
10. Check on other remarks sent by price collectors
11. Check on remarks sent through message module (new)
12. Perform price scaling according to editing rules[191]

[190] Periodical checks are performed on the sample of outlets and business layers; however these are not part of the monthly process and therefore are not discussed in the paper.

[191] Discounts and sales are divided into three categories: (i) regular discounts, (ii) options to receive or purchase at discounted prices other items and (iii) an option to receive more of the same product with a discounted price. According to the editing rules regular discounts are considered price reductions, options to receive other products are not price

The analysts are required to document cases that may not have been covered in past by the editing rules; these are discussed at the weekly meeting and a consistent solution is formulated to ensure identical treatment and minimization of bias in the editing process.

3.2.2 Monthly editing by CPI head office staff

After all transactions in the database have been edited by the commodity analysts, the computation of the product price indices and aggregate indices of consumption groups begins. These stages are observed by three types of statistical checks: (i) index levels (ii) index changes and (iii) contribution of indices to overall aggregates.

3.2.2.1 ***Index levels*** – Item price indices are aggregated into product indices. When products are homogeneous and/or services are subject to government controlled price regimes, deviations in item price index levels may indicate a flaw in the index. Therefore, at this stage, the outlying indices are detected by the system and double checked to ensure that the correct prices were recorded in the former stage.

3.2.2.2 ***Index changes*** – Price change boundaries are narrowed at the computation stage to allow for additional checks by the commodity analysts.
While logical checks in the price collection sub-system queried the price collectors when abnormal price changes were detected, the boundaries in the system must be quite wide in order to allow the price collector to complete the price collection in a fairly flexible manner. This is crucial to ensure cooperation of the outlet proprietors.

3.2.2.3 ***Index contributions*** – At this stage, contribution to the overall index is studied by the analysts for all five levels of the index, i.e. from the product index level up to the major consumption group level that comprises the total CPI. The contribution to the total CPI is a function of current relative importance (current weights) and actual price changes. Products or consumption groups that were major contributors to the total CPI are double checked to ensure that all indices are logical before publication.

The CPI analysts prepare a monthly report for each major consumption group and refer to any unique or abnormal activity in the index and present these at the monthly meeting held 48 hours

---

reductions and options to receive "more of the same" are considered price reductions in certain situations and require price scaling. These are double checked at this stage.

before publication of the CPI. At this meeting all groups are compared to ensure consistency throughout the CPI. The computer system was designed with standard queries from the CPI database to assist the analysts in their examination of the index. Highlights from the reports are reflected in the monthly press release and the monthly bulletin[192].

In section 3 we described several procedures of logical checks and process control that were enhanced by the introduction of modern technology in CPI compilation. In the following section we discuss some of the quality measures devised in an EFQM framework that were also enabled by the new computer system.

4. Quality assurance and EFQM

As mentioned, the computerization project of the CPI in Israel was performed under the assumption that automization will improve the quality of the CPI significantly. Intuitively, we feel that there has been an improvement in the quality of the CPI as a result of computerization. However, beyond intuition, statistical measures for a "quality improved" CPI were not designed during the computerization project.

Possible improvements in the CPI may relate to any of the four fundamental types of survey errors: *coverage, sampling, measurement and non-response*. The first two are considered as “after the fact procedures” and are treated with statistical analysis tools. The latter two are more concerned with procedural flaws during fieldwork or computation stages and may be treated with behavioral tools. By computerizing and elaborating work processes and methodology at all levels of the dissemination of the CPI; by building modern foundations and accessibility to data sources and by moving logical and statistical checks from the headquarters to the field, we believe that we reduced measurement and non response error and thus CPI quality has indeed improved.

In order to empirically measure the improvement in the CPI, we introduced, in November 2002, EFQM[193] methodology into the Consumer Prices Division. The advantage of EFQM is that it relies on input from the CPI staff, rather than assessing quality of the CPI by external auditors[194]. CPI staff participated in workshops for a period of six months and devised short-term and long-term

[192] The press release and monthly bulletin include all price indices computed by the Prices Department: CPI and output and input PPI's. The other price divisions review their indices in similar fashion to ensure consistency "across the board".

[193] EFOM is the European Foundation for Quality Management excellence model; a self-assessment diagnostic tool used to improve quality and performance. It is driven by result oriented philosophy and allows for flexible management strategy. For a detailed description see the Manual on PPI, pp. 318-319 or manual on CPI, pp. 224.

[194] External audits of a CPI are problematic due to the complexity and secrecy of the CPI.

projects to improve the quality of the CPI, and CPI surroundings including services rendered by the Statistical Office to the users of the index. Recommendations connected to issues such as database security, better dissemination to the public, etc. are beyond the scope of this paper and will not be discussed here.

As would be expected, many discussions during the workshops, focused on ways to estimate CPI quality before and after computerization. In this section we present several quality measures derived from these discussions in the EFQM framework.

It should be stated that a "true" comparison between price data collection using paper questionnaires and computerized data collection couldn't be conducted at this point, because of the complexity in comparing two different processes, and the fact that data from the pre-computerized era were not available[195].

4.1 Quality measures of the computerization

The assimilation of the automated price collection system into the CPI was staggered over an eight month period, between May – December 2001. During each stage, a group of price collectors moved from filling paper questionnaires to working with hand held computers (CAPI). The CATI was at the last stage. The data below starts at the beginning of the first stage, i.e. May 2001. The totals in the charts below are the number of observations existing in the database during each month. In the types of indices below we assume that changes in the following variables may indicate an improvement in the quality of the CPI:

1. Increase in number of "real" price observations
2. Decrease in number of observations with correction in the price made by the commodity analyst.
3. Decrease in number of observations with item remarks "temporarily ran out", "permanently ran out" and "unavailable".
4. Decrease in number of observations which an automatic query was raised with on line logical editing of data concerning the price.

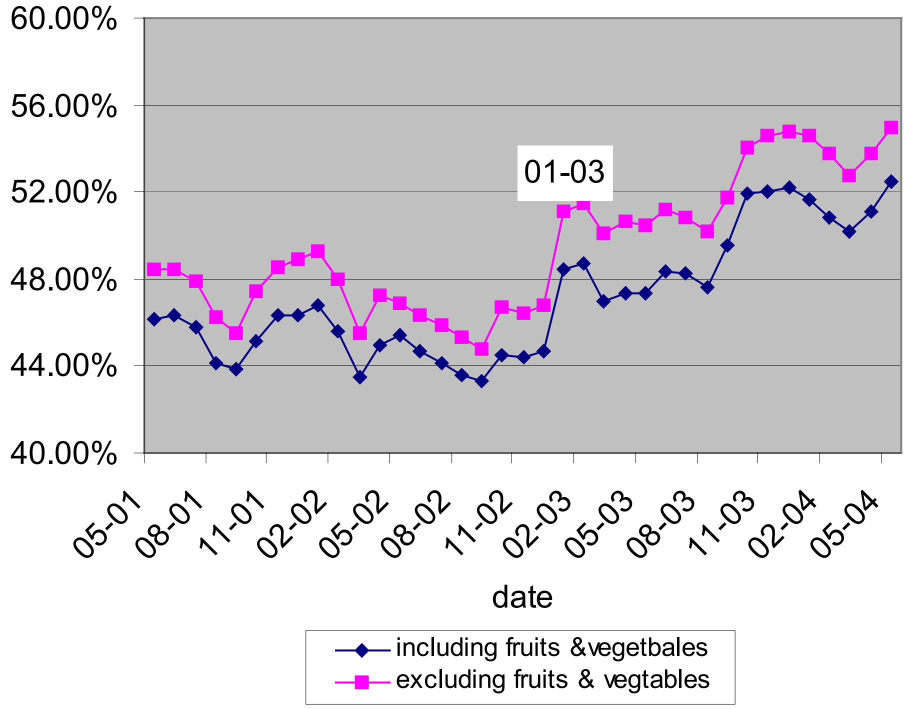
4.1.1 Number of real observations

Collecting prices using hand held computers vs. paper questionnaires allows better accuracy and transparency of the chosen product. In addition, the system leads the price collector in every

[195] Recently we have computerized the housing index and benefited from the EFQM process by designing several quality measures at the planning stage. These will be presented later on in this section.

step in a way that minimizes mistakes and a need to make a correction of the price by the commodity analyst. Thus, if the index were to improve over time, due to the computerized process, we would expect an increasing number of real prices compared to imputed ones.

Graph 1: percentage of real observations in the price collection sub-system

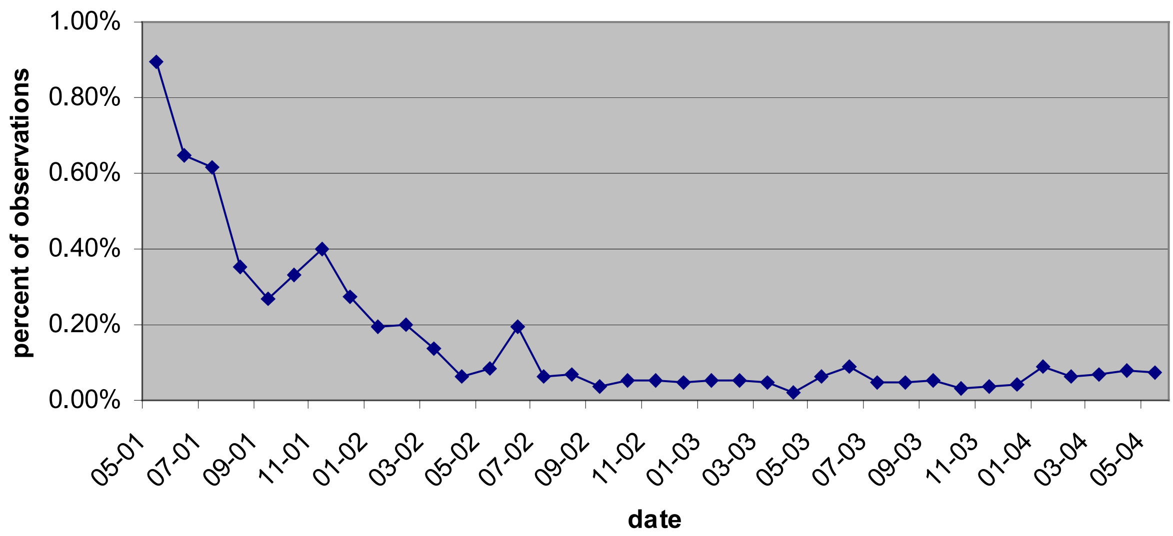


This variable "real observation" consists of situations in which a real price has been collected and thus there was no action of imputation. To be more specific, observations with item remarks "no change", "correction/completion", "quality/non quality change" and "new item" are included. Not included are observations with item remarks "temporarily/seasonally running out", "permanently ran out" and "unavailable". In graph 1 above we see an increase over time in the percentage of "real observations" since the beginning of the assimilation of the system – from 48.5% in May 2001 to 55% in May 2004. The percentage of real observations is always lower for the series including fruits and vegetables, due to seasonality imputations. Another interesting point is the decrease in real observation during the first eighteen months of the computerized system using the hand held computers. This may be explained by technical and other difficulties encountered by the price collectors in the early months of assimilation. After climbing along the "learning curve" the price collectors consistently increased the percentage of real observations from the last quarter of 2002 and until May 2004.

4.1.2 Number of observations with commodity analyst's correction of the prices

Due to the fact that the system leads the price collector during all the process of price collection, we expected to find a decrease in the mistakes made in the field and in the number of observations in which a price correction by the commodity analyst is required.

Graph 2: percentage of observations with corrections by commodity analysts



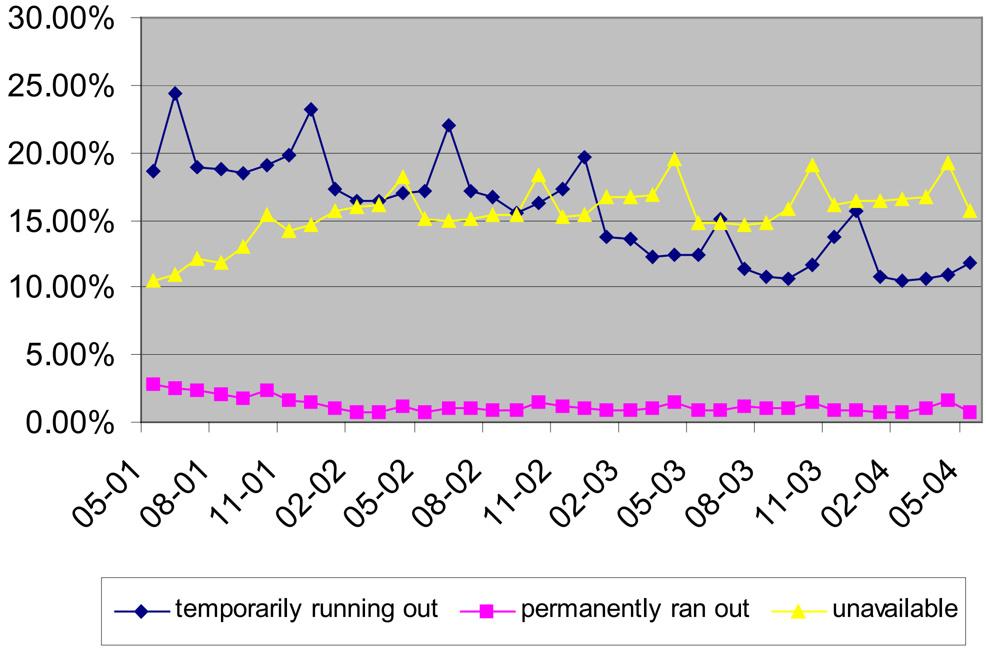
In graph 2 above we observe an immediate decline in the percentage of price corrections during the first stage of assimilation. The tendency of decline continues and after about one year into the automated price collection stability in the percentage of price corrections is achieved (less than one twentieth of a percent of the prices needed correction by the commodity analysts). Obviously, the price collectors have fully exploited the learning curve in this case.

4.1.3 Number of observations with item remarks “temporarily ran out”, “permanently ran out” and “unavailable”.

As we mentioned above (3.1.1.5), after three months that the chosen product was in status of “temporarily ran out”, a query is raised to remind the price collector to make sure that the sampled product is "back in business". If not, he should either replace the sampled product or change its status to “permanently ran out”. After three months that the product has been in the status of “permanently ran out”, a query is raised again to remind the price collector that he should either choose a new product for the defined item, or mark it as “unavailable”.
The instruction for handling those statuses were the same in paper questionnaires but it was much harder to enforce them. We found it interesting to see how the technological means influence the quality of fieldwork and check on one hand, if there has been a decrease in the

percentage of observations in statuses “temporarily/ permanently ran out and on the other hand check if there has been an increase in the percentage of observations in status “unavailable”.

Graph 3: percentage of observations with "non-price" statuses

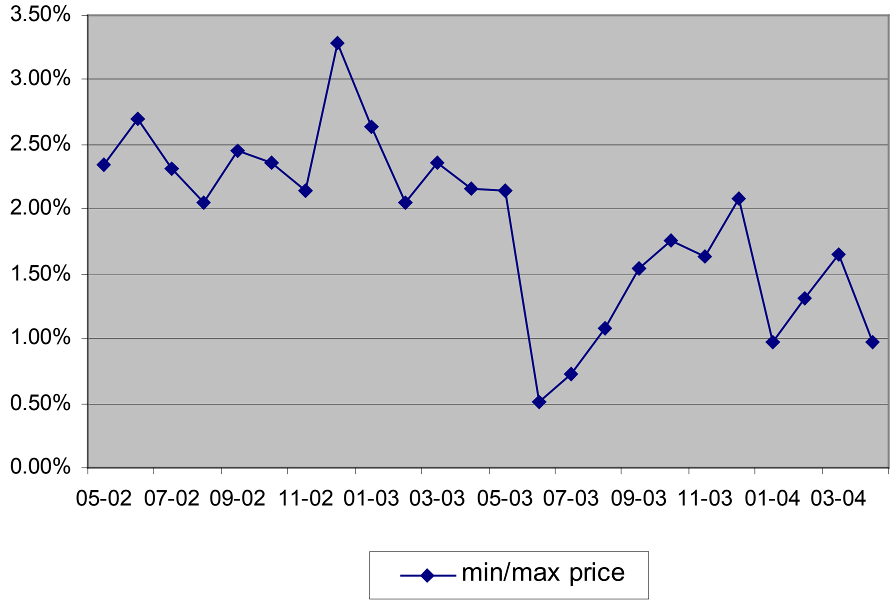


In graph 3 above we observe how the computer system has caused the price collectors to work more precisely according to the price collection rules. Over time the percentage of observations with temporarily or permanently ran out status has decreased (from about 20% to 10%) and unavailable status increased (from 10% to 15%. The increase in percentage of real prices, shown in graph 1 above is on account of the "non-price" statuses, therefore minimizing the need for price imputations.

4.1.4 on line logical editing of data

Although there are two kinds of on line logical editing of data in the price collection process - “percentage change” and “min/max”, we chose to relate only to min/max. On line logical editing of data concerning the percentage change, may reflect changes in pricing schemes or market conditions and not necessarily changes in quality of the CPI compilation. When comparing the percentage of observations outlying the min/max boundaries in the beginning of the computerized process with the present situation (graph 4 below) we can detect a significant change (from about 2.5% of price observations to about 1%).

Graph 4: percentage of observations outlying the min/max boundaries

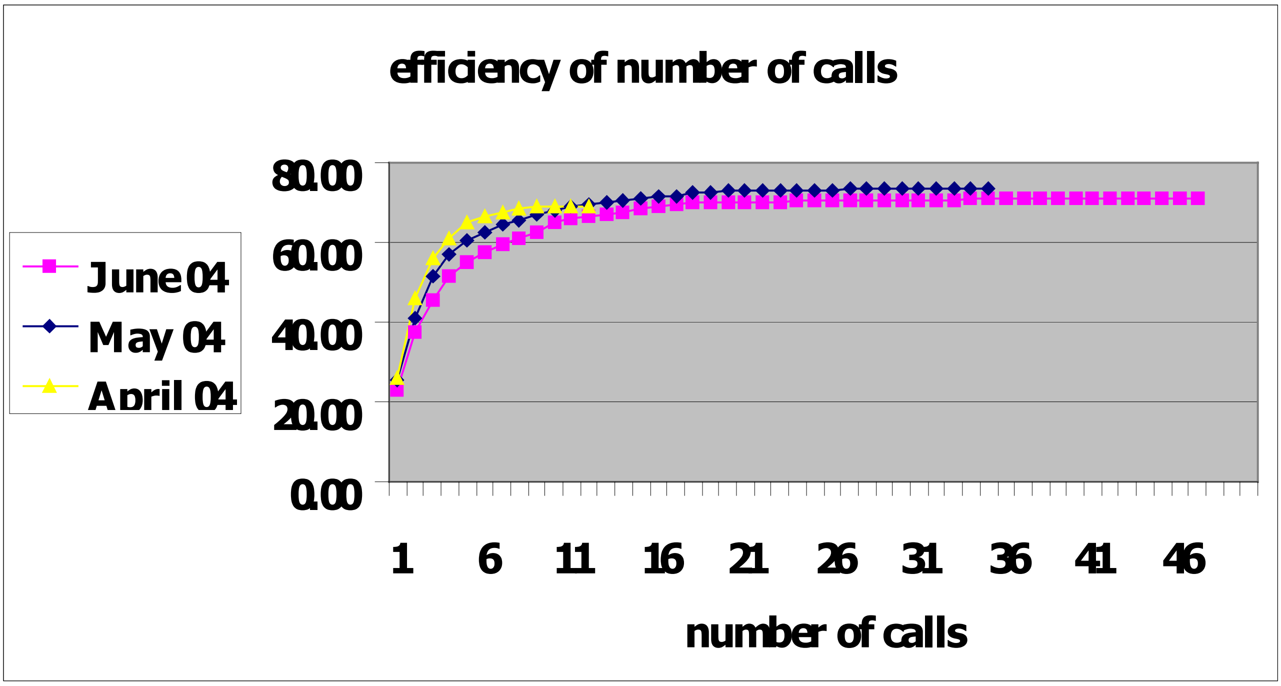


The improved procedures and control of the CPI collection process has decreased imputations and lowered variance, both at the price collection and index computation stages. Using this first set of quality measures has encouraged us that the computerization process has led to enhanced quality of the index, empirically and not only intuitively.

4.2 Price Indices of Housing

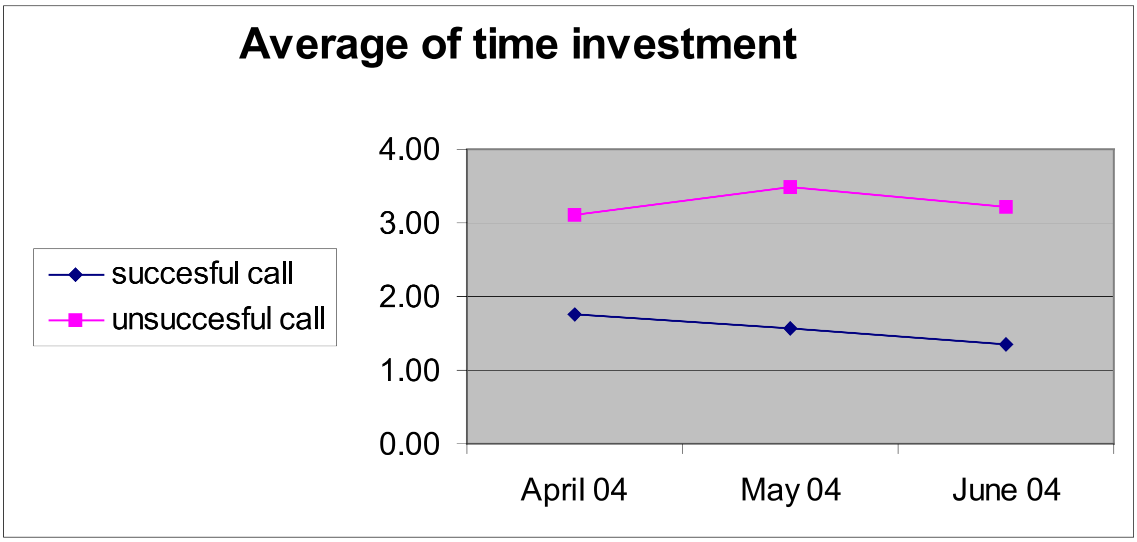
ICBS is currently (as of April 2004) assimilating a new module into the price collection and price compilation and analysis sub-systems. This module consists of the housing price indices, i.e. owner occupied housing and rent price indices. The fact that this module was being designed when the Prices Division was well into the EFQM workshops led to the introduction of possible quality measures, already during the planning stage. For example, computer assisted telephone interviewing allows for more efficient calling procedures and therefore enables the price collectors to attempt more calls for each sampled household. However, if one were to analyze the distribution of successful enumeration per number of calls, different decisions on what is an efficient breakeven point may be arrived at.

Graph 5: number of calls per successful enumeration of households in housing indices



In graph 5 we observe that 65% of the households were successfully enumerated by up till 10 phone calls; an additional 10 calls increased the percentage of completed households to 70% and any call above the 20th did not contribute to the index. Based on this kind of analysis one may decide (subject to budget constraints) what may be an optimal enumeration plan for the index. While we are only in the first months of the new module, by planning quality measures at the design stage and implementing them at assimilation stage, we hope to fine tune procedures in real time.
It should also be stated that the computer assisted telephone interview module for the rent indices leads to reduced costs per call, therefore enabling more calls per household; a recommended procedure only if it leads to higher response rates at the given budget constraints.

Graph 6: average time (in minutes) per calls in the housing index



In graph 6 above we observe a slight increase in the time invested for the unsuccessful calls[196]. In addition we see that the successful calls are handled more efficiently in the CATI module. Since unsuccessful calls can eventually lead to successful ones and increase the response rates of the household rent survey, we are now contemplating the possibility of allowing for increased calls per household and still remain within budget constraints.

The measurement tools that we implemented, although very humble at this stage of EFQM project, seem to indicate that there was an improvement in the quality of the CPI.

5. Summary and future work

ICBS completed a development project for the CPI that included enhancement of the technological systems along with improvement of procedures and methodology. Naturally, as in any technological project, most of the financial and human resources were invested in the development of complicated and modern technology and on training for the various users of the system. Quality measures concerning the influence of computerization on CPI procedures were not included as part of the design at the planning stage. Only after the officials felt that all systems were stable and resources were not expended all the time for problems connected to technological systems, we were able to address the issues of quality assurance and continuous improvement of the CPI. This is being achieved by the adoption of the EFQM framework for the CPI. This framework has led to many favorable results for CPI and its surroundings. In this paper we concentrated on measures for quality assurance and continuous improvement. These measures are centered on procedural errors in the index, and on minimization of non-response, i.e. behavioral tools. These in turn may reduce the need of statistical tools that address flaws in coverage and sampling.

Influence of computerization on the CPI is a difficult measurement task, especially when budget constraints do not allow for the operation of manual and computer systems in identical periods. Therefore the quality measures that we devised for the main components of the system can only indicate future improvements and not always explain what happened to the CPI during the "transfer period" from a paper based system to a computer based one. After the introduction of EFQM, additional modules that are being added on to the system, like housing price indices, are already benefiting from our building of quality measures at the design and assimilation stages. Cost-benefit analysis on procedural issues can lead to an improvement of the index in real time and still keep

[196] *Unsuccessful calls* are registered in the data base in one of the following cases: refusal, voice machine, non-connection, wrong number, etc. *Successful calls* are tallied only if the CATI questionnaire is completed as far as the

within the budget constraints. Future work includes expanding of quality measures for all fundamental types of CPI survey errors.

critical fields are concerned in the pre-determined design..

Rolf Boesch and Corinne Becker Vermeulen[197]: Quality assurance of the Swiss consumer price index

Kewords:
quality assurance (QA), total quality management (TQM), Shewhart-Deming cycle, process models, Swiss consumer price index (CPI), zero-error approach, outsourced decentralised price collection.

Summary:
An explicit quality management system (QMS) was introduced for the national consumer price index (CPI) following a breakdown in the 2000 revision. The CPI total quality management system (CPI-TQM) is based on a zero-error approach combined with a process model.
In practice, the relevant results are identified and subjected to the necessary comprehensive scrutiny. For this, processes are analysed using a model. In this connection, the risk structures and the adaptation of implementation methods vary depending on whether the processes are on-going, such as the production of a price index, or rare or even one-off events, such as index revisions.

Introduction

Index breakdown despite all precautions

Since the start, quality assurance has been a main principle underpinning the Swiss national consumer price index (CPI). It is of course standard practice that each result is checked by a second person and that the plausibility of each monthly index is justified.

Nevertheless, there was a breakdown when the methodology of the CPI was revised in 2000. In the process of underweighting energy sources, units of energy [joules] were used instead of expenditure [CHF] and the fact that the price per unit of energy [CHF/joule] is different for electricity, gas and oil was overlooked. Furthermore, the person who checked the figures did not notice that the energy unit prices had not been specially adjusted. Despite this mistake, the overall relative weighting seemed to be quite plausible and the variation since the last weighting was carried out in 1993 did not seem impossible. This type of mistake – an omission or an apparently plausible but incorrect thought process – is much more difficult to detect than an item placed in the wrong category or an incorrect calculation.

When the price of oil rose dramatically in late summer 2000[198] public interest focused on the weighting for this factor and after a few weeks the mistake was discovered[199]. There was a widespread reaction among the general public. This confirmed the fact that reliability of the CPI is extremely important to the Swiss population since the CPI is used as a compensation index. It represents a neutral yardstick that is used as a basis for salary negotiations between social partners and is an important factor for calculating pension adjustments and for modifying contracts. Furthermore, the Swiss National Bank (SNB) uses the CPI to evaluate its inflation objectives.

---

[197] Swiss Federal Statistical Office
[198] SFSO: Producer and import price index for September 2000, Neuchâtel October 2000 (only available in French and Gernan – www.statistik.admin.ch)
[199] SFSO: National consumer price index for November 2000, Neuchâtel November 2000 (only available in French and Gernan – www.statistik.admin.ch)

One of the measures taken by the Swiss Federal Statistical Office (SFSO) as a consequence was to introduce a quality management system for the national consumer price index.

CPI Total Quality Management: Basic principles of the CPI-TQM

Zero-error approach

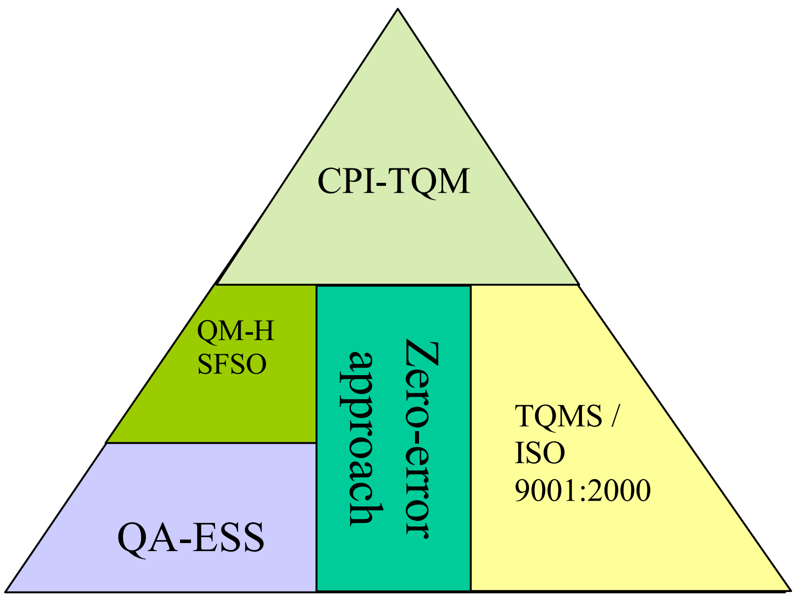
As clearly stated in the introduction, the aim of the CPI quality management system (QMS) is to avoid having to revise CPI results. This automatically means that there can be no room for error.

Errors can of course never be completely ruled out. In contrast to many classical QMSs in the manufacturing sector, in the case of the CPI there is no leeway with regard to the precision of the published results. Likewise, there is no room for error concerning the results of the relevant activities, all of which are subject to verification. In general, QMSs that include and regulate all processes are known as total quality management systems (TQMS). In addition “total quality” corresponds to the “zero-error approach”. Including both requirements, we therefore refer to the system applied with regard to CPI as CPI total quality management (CPI-TQM).

Development of CPI-TQM from general QMSs

In the general sense, ”quality” is basically defined by the recommendations of the European Statistical System (ESS), which is also the basis for the SFSO’s general Quality Management Handbook (QM-H). (This handbook does not exclude result revisions, however). The corresponding quality requirements for public statistics[200] are regularly reviewed in relation to the CPI.

The Shewhart-Deming cycle discussed below forms part of both the SFSO’s QM-H and the Q-norm ISO 9001:2000.



[200] According to the SFSO’s QM-H these are: relevance, transparency, reliability, comparability, impartiality, security and availability.

Figure 1: CPI-TQM is based on general standards applied within the SFSO and elsewhere

QMS tools should be used by the qualified personnel directly involved

Quality assurance tools should be made available to the existing organisation of specialised personnel and line managers, and staff should be trained to use them according to their level and given the necessary support. The line manager checks process descriptions and ensures coherency. Quality assurance determines how tools should be used, ensures training and support and verifies the use of the tools at regular intervals.

Applying CPI-TQM

Analysing processes as a starting point for the practical application of a QMS

Two questions should be asked as the outset:

- What level of detail should be monitored?
- How intense should controls be?

In order to answer these questions all processes need to be thoroughly analysed and the results should be set out in a coherent overview.

What level of detail should be monitored?

In CPI-TQM, monitoring should focus on including all processes that produce results at the lowest level of actual application.
All quantitative results and intermediate results that are relevant for the published results of the CPI should be included and checked according to a defined method.

For building up the process hierarchy (figure 2), the details of the organisational links at a higher level are of secondary importance and should be determined as simply and quickly as possible. For example, the existing organisational structure or the archiving structure can be used (in this way no new additional structures are created which could lead to inconsistencies).

What is important, however, is the flow of intermediate results right through to the published result. This must be included in its entirety.

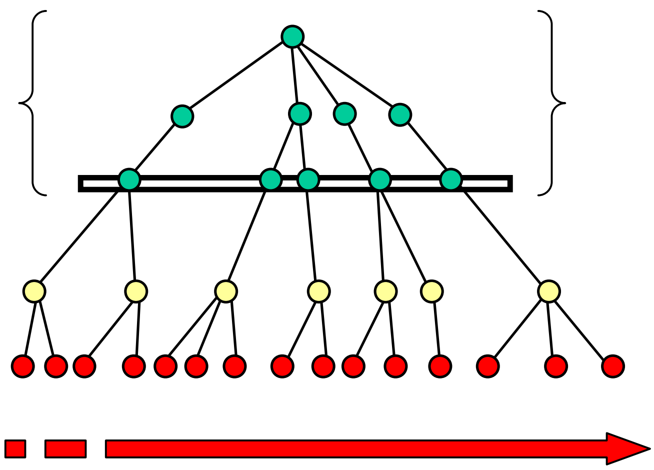


Figure 2: Practical diagram of the process hierarchy with results (dark / red circles). The important aspect here is that they are included right through to publication of the result.

How intensive should the controls be?

This question can only be answered according to the individual circumstances and should also be monitored at regular intervals. The results of a risk analysis of the process reveal what kind of errors may occur and how they can affect the result of the CPI. Subsequently, the standard process to be used as the model for the analysis should be flexible enough in its parameters that the necessary controls can be defined.

During the development and practical application of CPI-TQM we identified the 3 design parameters shown below. A sub-model has been devised for the density of the controls with up to 4 different levels of controls. This model is described below.

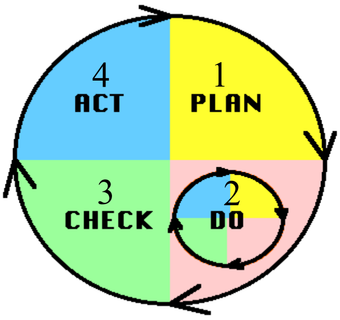
Design parameter	Question	Remarks
Frequency of controls	How often is a process divided into	The aim here is to identify all quantitative results and all decisions that could affect the published

	individual activities and the partial result at their "interface" controlled?	result of the CPI, and to include them in the controls.
Density of controls	How many people check the partial result at the "interface"?	The CPI-TQM model includes the following 4 levels (roles) that can be used depending on the importance of the result: - SELF: self-checking (compulsory) - PEER: specialist within the team who was not involved - EXT: specialist from outside the team or the office, in particular for the plausibility aspect - APPROVAL: approval of the result with checks by the responsible superior (compulsory)
Intensity of controls	How intensive should the controls be?	Description of the necessary examination method. In the 2005 revision of the CPI, the following standard control methods have been defined so far: - Cross-checking - Check list - Internal plausibility check - External plausibility check - Calculation check - Parallel construction (the result is built up twice separately from the basic figures). In addition, there is a growing collection of check lists available.

Table 1: Three design parameters to be used for setting out the necessary and adequate control process

Available process models

As in the case of most modern QMSs, in CPI-TQM each process is constituted as a cycle of rules. The "plan, do, check, act" cycle (known as the Shewhart-Deming cycle)[201],[202], which must be repeatedly applied, ensures that requirements and results are continually compared for evaluation and monitoring purposes, so that the necessary measures can be identified and taken[203].



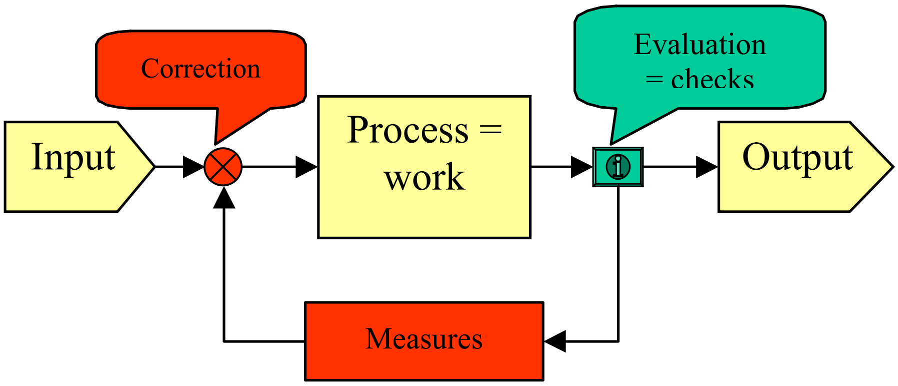


Figure 3: Two diagrams showing a regulated process: top, according to Shewhart and Deming; bottom, closed-loop control system (as used in automation)

Regarding Figure 3, the fact that the outer cycle encompasses the inner cycle in the top diagram should be mentioned first of all. Frequently, the plausibility of the aggregated results also has to be controlled. The process description itself is also a result that should be checked at regular intervals.

Despite the fact that the Shewhart-Deming cycle is easy to show in diagram form and to comprehend, it is still purely theoretical. A more suitable version for analysing the process is the diagram of the closed-loop control system with its variables and parameters for automatic control (bottom, figure 3). This version, however, is based on an approach where process errors are

201 As part of his research into statistical process control (SPC), Walter A. Shewhart developed the "plan, do, check, act" cycle in the Bell Laboratories in the 1930s. His colleague W. Edwards Deming introduced this cycle in the industrial sector from the 1950s onwards, achieving his early success in Japan (TQM originally stood for Toyota quality management). This success was then repeated by Deming at the beginning of the 1980s, starting with the famous white paper entitled "If Japan can, why can't we?", published in the USA by NBC.

202 The PDCA cycle devised by Shewhart and Deming is also part of the ISO 9001: 2000 Q-standard (Section 0.2).

203 In the natural sciences and technology sectors the principle described here is known as feedback. Requirements and results are continually compared and the necessary adjustments are then calculated. In recent decades, the electronic implementation of (closed-loop) control systems in automated processes in particular has gained enormous importance in relation to quantity and quality of industrial output.

continually eliminated by applying a compensatory function at the entry point whenever necessary (integral calculation similar to an analogue computer).

The following procedure model for CPI-TQM therefore combines the desirable characteristics of both basic models.

The definitive process model used for the Swiss consumer price index

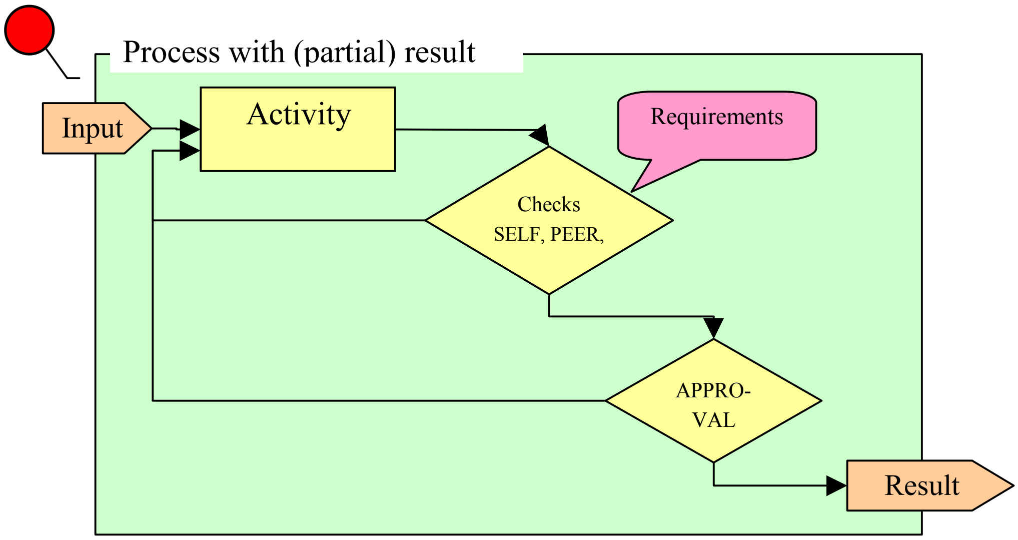


Figure 4: An individual process (activity) in CPI-TQM with its main characteristics

Parameter	Contents, definition
Input data	Are they quality-assured?
Result (Output)	Description
Activity	Description of aims, processes, methods
Requirements	What criteria must be met by the result?
Checks	SELF: who, when, how (compulsory)
	PEER: who, when, how
	EXT: who, when, how
	APPROVAL: who, when, (how) (compulsory)

Table 2: Characteristics of a CPI-TQM activity

The model of the individual process includes the parameters that are to be evaluated by analysis and the characteristics that are used to define the controls. These can then be easily defined and set out in such a table.

For CPI-TQM the tables are drawn up in a simple database which enables fixed data (names, functions) and various formats of the same content to be used over again for different people. The result is that those responsible for individual tasks have a simplified overall diagram and one page per allotted task, while the line manager has a complete catalogue with all the details of all the tasks.

Practical application of CPI-TQM in relation to CPI

Process security in the production of the CPI

Together with the staff, being the primary sources and users of know-how, we analysed the activities from the point of view of process characteristics (see Table 2) and stored the results in the database using brief descriptions. In addition to the elements that describe the process, **information about documentation** is also included: in many cases, there were already detailed working instructions in the production catalogue. These were checked and included through a link.

Defining the controls was of primary importance: results are never obtained without cross-checking, but it is not always necessary that checks be carried out at all 4 levels (SELF, PEER, EXTERNAL, APPROVAL; see Table 2). As a rule of thumb calculations require only an intensive peer check plus a plausibility check, which in most cases entails a recalculation of the most important contributions (internal plausibility check). For methodological decisions regarding concepts, external experts are called upon (for example, from the Evaluation and Methods Group) and comparisons with other concepts or data are made (external plausibility check)[204].

Finally, the process fact sheet for the task was passed on to the person responsible for application and updating (in the case of modifications, the approval of the line manager must be sought and the quality manager should be informed).

As indicated in the above representation of the Shewhart-Deming cycle, a distinction can be made between inner, direct control cycles and outer control cycles involving several processes (the latter also include the regular evaluation of the actual process descriptions). Some examples are listed below.

Each month, the staff responsible for producing the CPI hold a quality assessment meeting between the preliminary calculations for the new index and publication. At this meeting, a structured assessment is made of that month's index production. The problems encountered during price collection, divided into centralised and decentralised price collection, are discussed. Any special requirements in the coming month are noted. A report is drawn up which includes the necessary measures.

Before the index is approved for publication, a number of line and staff experts study the monthly press release as well as checking its plausibility.

Special case: Outsourced, decentralised price collection

In the 2000 CPI revision, special emphasis was placed on the system for decentralised price collection. The prices of all products with evident regional differences were collected in 24 regions. The quality of the prices noted as entry data for calculating the index constituted an important element in relation to the quality of the index.

Before the 2000 revision, 24 local authorities were responsible for collecting prices. This meant that the SFSO was in contact with 24 different local authority offices, which each had their own ideas about collecting prices. Consequently, in some cases the SFSO's task of monitoring and implementing a standardised price collection system was difficult, to say the least.

[204] A comparison often used in the CPI is based on the Eurostat HVPI, which is generally considered to be a good summary of current methods and definitions.

From the point of view of CPI-TQM, this meant that the requirements concerning price collection could not be set out in a uniform way, neither could the collection process be sufficiently implemented and controlled. This fact was all the more important since the problems and methods involved in calculating a price index today, such as variations in product ranges and changes in quality, are becoming increasingly demanding. An improvement in the price collection system was therefore urgently needed.

A good solution was found in that decentralised price collection was outsourced to a professional sub-contractor from the market research sector[205]. Here the compulsory methods and requirements can be set out in a price collection handbook. This must also be accompanied by suitable training and the implementation of the requirements (checks, feedback and improvements).

As far as concerns montly price collection, controls are of great importance. The market research institute constitutes an inner cycle: it has to control the requirements regarding the results itself. The results achieved by the market research institute regarding quality assurance are summarised in monthly reports that contain a large number of figures and indicators, for example the prices collected per collector, the number of products replaced and a statement concerning the controls made and the corrections carried out by the institute.

This report already enables the SFSO to satisfactorily monitor decentralised price collection. In a broader cycle, direct checks are carried out by the SFSO and the plausibility of the information is verified a second time. For this, as for centralised price collection, the SFSO team has the back-up of the database for price statistics: all price changes outside certain limits are highlighted. At the same time, a list of all changes in products and ranges is compiled.

These plausibility lists are checked item by item. Telephone enquiries are made at sales points and on-the-spot checks are carried out. In addition, price collectors are monitored at regular intervals or the prices collected are rechecked. An important element here is notification by the SFSO of the errors it finds. Every month the quality of the price collection system is discussed at the SFSO and sent to the market research institute int he form of a quality report.
This is a practical example of how the cycle works: for a while, the same errors made by the decentralised price collectors had to be corrected every month. Finally, enquiries revealed that the market research institute had not passed on the remarks concerning quality made by the survey manager at the SFSO to the price collectors in an appropriate form. The “act” element in the Shewhart-Deming cycle was being omitted, i.e. the information was not being passed back to the price collectors working in the field! After this problem had been eliminated the errors disappeared almost completely within a few months.

The SFSO retains “ownership“ of this whole process of price collection. Intensive controls, including decentralised price collection, and a strong presence at the 3 training sessions for collectors organised every year are as important as revising and improving the price collection handbook. The necessary level of excellence can only be achieved through maintaining intensive two-way contact and a full-hearted commitment on both sides: the sub-contractor works according to market conditions, but with a given and monitored level of quality.

Process security in the annual weighting of the CPI

[205] Corinne Becker Vermeulen, “Subcontracting the price collection: the swiss experience”, paper presented at the Joint ECE/ILO Meeting on Consumer Price Indices (Geneva, 1-2 November 2001) (available at: http://www.unece.org/stats/documents/2001.11.cpi.htm).

Since the 2000 revision, the annual weighting of the CPI basket of goods and services of the year *t* has been calculated from the results of the Swiss Household Budget Survey (HBS) carried out two years previously (HBS *t*-2). Since for historical reasons the 5-digit levels of the nomenclature used for the two sets of statistics do not show a 100% match and some categories have to be carefully split according to alternative formulae (top-down principle), certain measures are necessary.

Like the on-going calculation of the CPI, the entire process was analysed and documented using the CPI-TQM model. In contrast to on-going price surveys, the risk with the weighting system is greater, however; on the one hand, the weightings have a direct influence on the whole CPI for one year and later on the interlinked Laspeyres chain index, and on the other there is a certain amount of pressure to meet the deadline because the data obtained from the HBS are finally available only in the last quarter before the adoption of the new CPI weightings.

The consequence is that plans have to be drawn up with great care and the 4 control levels are used to full capacity. In addition, the entire weighting system is checked several times; the plausibility of the entry data from the HBS is verified using market data obtained in advance and a comparison with the previous year. The final basket of goods and services weighting is also compared with the weighting for the previous year, price trends in the different categories of items and market data.

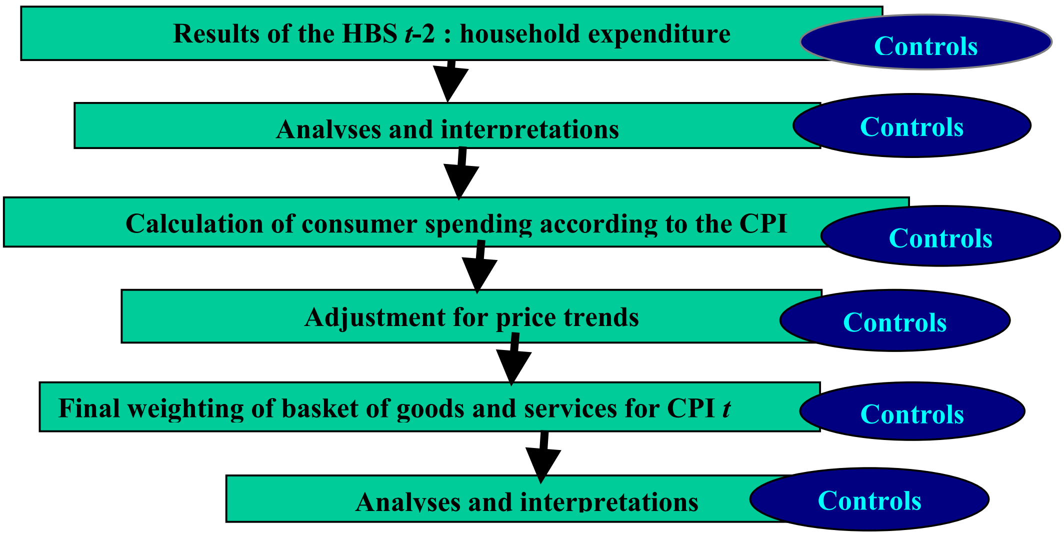


Figure 5: Transformation of the SSIC data for CPI weighting with checks at every stage

Finally, the weighting system is presented to principal CPI users at various workshops. Those attending these workshops include representatives of the regional statistical offices, social partners and forecasters from various sectors ranging from economic research and the Swiss National Bank to commercial banks and universities.

In this way, any possible errors are almost certain to be discovered. A further advantage of the workshops is that, thanks to the close contact with researchers and users of the CPI, there is always useful feedback concerning users' needs, while at the same time the users have an opportunity to learn more about the possibilities and limits of drawing up a price index.

Quality assurance in the 2005 revision

The CPI is revised every 5 years. Each revision project, which is in fact unique, involves a major risk of a breakdown: in any one such project it is far more difficult to ensure that all relevant results are of the highest quality and in particular contain absolutely no errors[206] than in the monthly production of the CPI or in the annual weighting. It is only partly possible to repeat processes that have already been planned and evaluated!

There are two further typical risks with the project: the risk of setting a wrong goal for the project and the risk of not meeting the deadline. The date when the revised CPI will be published is announced well in advance and cannot be changed.

SFSO statistical project model

Accordingly, the SFSO has drawn up a project management handbook (PMH)[207] that contains a compulsory model for statistical projects.

The PMH indicates the first effective measure for reducing the risk of setting a wrong goal: users' needs should be analysed at the start of the revision process[208].

In order to further reduce both risks, the statistical project should be divided up into various phases. When each phase is completed, a corresponding report is drawn up and submitted to various bodies up to and including the central management of the SFSO, so that despite the uniqueness of the project there is internal repetition and the Shewhart-Deming cycle can be applied.

The SFSO project model also includes a rough outline of the results expected from a statistical project; this outline can be found in Appendix I. It constitutes the upper part of the process hierarchy (see Figure 2) for the 2005 CPI revision project. Further down the hierarchy, the concrete tasks comprising the project are identified by the project manager and allotted to the staff. The latter may then divide their tasks into sub-tasks. The yardstick for dividing tasks is the relevant result, however, that may influence the published CPI for 2005 and must therefore be controlled.

On-going analysis of processes under the project conditions

For the CPI revision, the scientific staff form sub-teams for each task (matrix organisation). This ensures internal exchange and feedback in the sub-team that helps them to work in a targeted and efficient way. This internal exchange is well established in the whole CPI team and is further encouraged for the revision process by weekly team meetings where those present are systematically informed of the progress being made and specific technical problems can be discussed by the team as a whole.

In the on-going processes, fixed job descriptions are enhanced by outlines and control plans for each task that the sub-teams work on independently. As the work advances, the outlines become working reports that are regularly checked by and discussed with the project management. At the end, these working reports, which are much more detailed than the general phase reports, set out decisions and procedures, as well as unsolved problems and unanswered questions that go beyond the project.

The control plans that the sub-teams draw up independently according to the prerequisites of the CPI-TQM process model (a form is provided for this purpose that indicates the necessary variables and parameters as well as instructions for the analysis) are checked by the CPI change manager and discussed with the sub-teams. At the start of a phase in the process, it is important to recognise whether a task should produce quantitative results (calculations) or qualitative results (conclusions,

[206] The zero-error approach remains an absolute priority with regard to the published results of the CPI.
[207] See PMH V5.3
[208] A survey was carried out among IPC users in autumn 2003.

concepts, proposals) (see Section 4.1 concerning the use of the 4 checking levels). The amount of work involved in the checking and the staff that will be needed have to be estimated. Schedules then have to be drawn up with peer colleagues and other controllers.

All the controls made are described in the control plan, which thus becomes a control report. Check-lists from the SFSO project handbook are also available for carrying out the controls. In the case of particularly difficult tasks, special check-lists may be devised specifically for the project and approved separately. The change manager re-checks the final reports from the point of view of quality assurance and the project manager uses them for final approval.

The systematic planning described here and the accompanying documentation involves additional work, even when the system is optimally integrated. To the individual person writing a report, the information collected may appear at that moment to be trivial. Furthermore, awareness that the planned schedule can never be achieved 100% may cause uneasiness. The processes of collation, evaluation and synthesis, however, are prerequisites as well as the price that has to be paid for ensuring that measures can be taken in good time to guarantee that high-quality content is achieved and the deadline is met.

Since both the SFSO project model and the CPI-TQM for the 2005 revision are at an introductory stage and quality assurance is in general still being set up in the service sector, both enjoy the support of specialised consultants. When the consultants were selected, emphasis was placed on their proven success in complex projects (in fact in IT and organisational projects) and not only on their theoretical knowledge of project and quality management.

General check links

The SFSO project model stipulates that the results of the different phases of the project are to be set out in a report which is then discussed in various stages and is approved and passed on through the SFSO lines up to general management.

The first general link represents audits carried out by the project principal, who must be convinced that, in each phase, the approach adopted corresponds to the given aims and requirements and that suitable methods and controlling procedures are being used.

First of all, the results of each phase are submitted to a supporting panel of external experts[209] who give their opinion. Their comments and proposed measures form an integral part of the phase report and of the proposal that approval be given by general management for the next phase.

The phase reports are then distributed to all interested parties. This includes on the one hand all the users already mentioned, but the bodies involved in the price collections, representing various sectors, are also important. The entire composition of the basket of goods and services and the price collection apparatus is discussed in detail at numerous meetings with representatives of the various sectors, thus ensuring the best result.

Conclusions drawn from the QMS experience vis-à-vis the CPI

A decisive factor in the success of a QMS is the integration of Q-tools into the processes. These tools must be made available to the lines and those working in the field, who should be appropriately trained to use them and regularly monitored. Furthermore, QMSs should not be allowed to have a life of their own. Quality assurance must not be an imposed, foreign element that

[209] Generally, the same as those invited to the annual workshops after the CPI basket of goods and services has been reweighted (see 4.2)

uses up additional resources but remains an outside factor (for example, as an appended, additional organisation). If the QMS is not properly integrated many errors will remain undetected, those with knowledge and experience will be incapacitated and the responsibilities they take on will be reduced.

If a QMS is to be really well integrated into the processes it must be simple and efficient, while at the same time a QMS that is easy to use and always up-to-date requires the type of electronic support that is not yet available in a simple, integrated form. With regard to the CPI, the CPI-TQM is also still at the construction stage and is evaluated and improved in every phase with the help of specialised consultants.

A systematic QMS involves additional work. This is the price that has to be paid for ensuring that quality assurance measures will be identified without fail and in good time. If in practice there are difficulties meeting the deadline, these measures mostly consist of setting priorities and deciding on which tasks to abandon, since no additional resources are available.

In the case of decentralised price collection where the SFSO implemented the change of system already mentioned for the 2000 CPI revision and mandated a market research institute to collect prices, experience shows that the decisive factors are: precise specifications concerning the services required, continual and detailed checking of the results received and on-going optimisation of the collaborative aspect.

With unique projects such as the revision of a price index, a phase model can serve to ensure artificial repetition on the one hand, and risks can be reduced through proven structural prerequisites, forms and check-lists on the other. This is achieved through the project management handbook that is being introduced at the SFSO.

It is also possible to improve security with regard to content and deadlines within each phase through repeated checks, for example through regular presentations, reports and discussions of the partial results. Existing process documentation in continually repeated processes is replaced by “on the fly”, on-going process analyses by those responsible for individual tasks, the analyses then becoming work and control reports.

For important results all possible means of controls, including the double calculation of important results, must be fully used. It is also important that the methodological results of revisions are accepted by internal and external experts and users before they are implemented.

Appendix I: Sections and results of the SFSO project model with 5 phases and chapter titles

Sections and results	**Phases**				
	Start	**General concept**	**Detailed concept**	**Realisation**	**Introduction**
X10.0 Purpose	**110.0**	**210.0**	**310.0**	**410.0**	**510.0**
X20.0 Procedure	**120.0**	**220.0**	**320.0**	**420.0**	**520.0**
X30.0 Results	**130.0**	**230.0**	**330.0**	**430.0**	**530.0**
X31.0 Project management	***131.0***	***231.0***	***331.0***	***431.0***	***531.0***
X31.1 Project plan	131.1	231.1	331.1	431.1	531.1
X31.2 Project report		231.2	331.2	431.2	531.2
X31.3 Project administration		231.3	331.3	431.3	531.3
X31.4 – X31.5 Phase results	131.4	231.4 / 5	331.4	431.4	531.4
X32.0 Project realisation		***232.0***	***332.0***	***432.0***	***532.0***
X32.1 Basic conditions		232.1	332.1	432.1	
X32.2 Methodological framework		232.2	332.2	432.2	
X32.3 Organisation and processes		232.3	332.3	432.3	
X32.4 Resources		232.4	332.4	432.4	
X32.5 IT		232.5	332.5	432.5	
X32.6 Tests		232.6	332.6	432.6	
X32.7 Training			332.7	432.7	532.7
X32.8 Communication		232.8	332.8	432.8	532.8
X32.9 Overall system		232.9	332.9		532.9
X33.0 Quality assurance		***233.0***	***333.0***	***433.0***	***533.0***
X33.1 Checking plan		233.1	333.1	433.1	533.1
X33.2 Check on completeness and deadline		233.2	333.2	433.2	533.2
X32.3 Check on content		233.3	333.3	433.3	533.3
X40.0 Consultations and decisions	**140.0**	**240.0**	**340.0**	**440.0**	**540.0**

Session 5: Other Issues

Moderator: John Greenlees, BLS

Summary of Session

This session comprised three papers, all applications of the hedonic regression approach to quality adjustment in high-technology product areas. The presentations and discussions were more concerned with broad issues in hedonic methodology than with the regressions results themselves.

The first paper, by Mick Silver and Saeed Heravi, used data on desktop computers to analyze the differences between time dummy hedonic (DTH) indexes and hedonic imputation (HI) indexes. Using the Törnqvist specification, the authors decomposed the difference between the two methods into two factors corresponding to (i) parameter instability and (ii) changes in average product characteristics between the two periods being compared. They concluded that HI indexes are preferred when parameter instability dominates, and DTH indexes when quality characteristics are changing. When both sorts of changes are occurring, it may be best to average the index estimates from the two approaches.

The second paper was by Jan de Haan. Stimulated by Silver and Heravi's results, de Haan also analyzed the properties of time dummy hedonic indexes using data on PCs. He noted that the time dummy hedonic index is a special case of the hedonic imputation index, but only if the regression weights are properly chosen.

The final paper in this session was a comparison of different quality adjustment methods for digital cameras, by Kari Manninen. The author concluded that the hedonic method is feasible and sufficient, even in situations where quality change is rapid. He also noted that simpler and more complex specifications produced similar index results. Thus, his results suggest that the data requirements for a hedonic index may not exceed those necessary to apply the matched-model approach.

In leading off the general discussion of the three papers, the Chair noted that several issues had been raised, among them the acceptability of the time dummy approach, the appropriate set of weights to be used in hedonic regressions, the method in which imputation approaches are employed, and the adequacy of the matched-model approach.

Several participants argued for more research on the extent to which the DTH and other hedonic models are robust to changes in preferences, missing variables, and other factors. Martin Ribe noted that research at the European Hedonic Center found much wider differences across specification than Manninen's paper found. There was broad agreement on the desirability of cross-validation work. Mick Silver noted that the greater the spread between different index results, the less confidence we can have in either.

David Johnson observed that in order to evaluate whether these approaches were useful in operational price index work there was a need to focus on month-to-month changes, not just long-run indexes. He also noted that both bias and variance were issues in determining which approaches were "better."

In response to a comment by Kevin Fox, Jan de Haan noted that weighting could be justified by heteroscedasticity considerations as well as a desire for representativity. He emphasized that equal weighting in a hedonic regression is difficult to reconcile with standard index formulas.

Two participants argued that multi-period DTH models were not operationally feasible and that our analysis of the time dummy approach must therefore focus on two-period models.

Recommendations for Statistical Agencies

1. There was agreement that weights should be used in regressions where possible. Furthermore, with respect to DTH models, there was agreement that de Haan's criticism of Silver and Heravi's weights was accurate.
2. The issues surrounding hedonic regression are how they should be employed, not whether they are acceptable or desirable. There was little or no support for the argument sometimes made that the matched model approach is sufficient when high-frequency data are available.

Mick Silver and Saeed Heravi: The difference between hedonic imputation indexes And time dummy hedonic indexes for desktop pcs

Acknowledgements are due to the U.K Office for National Statistics (ONS) for supplying the data for this study. This follows a well established research link with the ONS which has been of much benefit to the authors. Naturally the ONS are not responsible for any views, errors of omission or commission in this paper. We acknowledge useful comments from Ernst Berndt (MIT), W. Erwin Diewert (University of British Columbia), Kevin Fox (UNSW), Jan de Haan (Statistics Netherlands) and other participants at an NBER/CRIW conference on index number and productivity measurement in Vancouver, June 2004.

ABSTRACT

This paper shows why the two main approaches to estimating hedonic indexes can produce quite different results. The proper measurement of inflation requires the use of hedonic indexes, rather than matched models, for product areas which have a high turnover of differentiated models. The two main approaches to hedonic indexes are *hedonic imputation* (HI) *indexes* and *dummy time hedonic* (DTH) indexes. HI indexes value a fixed period's basket of characteristics using both base period and current period hedonic coefficients and take the ratio of the latter to the former. HI index number formulas differ in their use of which period's characteristics are held constant for the valuation. DTH indexes estimate price change using the coefficient on a dummy variable for time in a hedonic regression which uses both base *and* current period's data. For DTH indexes the slope parameters are constrained to be the same for both periods to allow the intercept shift to measure quality-adjusted price change. For HI indexes the change in the parameters over time are, paradoxically, the essence of the measure.

The study provides a formal analysis of the difference between the two approaches. It shows the conditions under which the approaches will provide similar results which, surprisingly, may even be when parameters are unstable, and the factors governing differences between the results. It shows that differences between the methods can be substantial and discusses why this is the case and the issue of choice between these measures. An illustrative study for desktop PCs is provided. It demonstrates the importance of using hedonic indexes for product areas which have a high turnover of differentiated models, that the two approaches can differ and the factors underlying the difference.

KEYWORDS: HEDONIC REGRESSIONS; HEDONIC INDEXES; TIME DUMMY VARIABLE INDEX; CONSUMER PRICE INDEXES; PERSONAL COMPUTERS

JEL classification: C43, C82, E31.

The purpose of this paper is to outline and compare the two[210] main and quite distinct approaches

[210] An alternative approach to hedonic indexes is to use a fixed effects panel estimator in which instead of hedonic characteristics on the right-hand-side of the regression, each model, other than a benchmark one, has a dummy variable (suggested by Diewert in (2002a)—see Aizcorbe (2003) for an application). In practice the panel estimation procedure is much simpler; the price deviations for each model from their mean over time are regressed on the deviations of the explanatory dummy variables for each model for all *t* with an adjustment for degrees of freedom (Davidson and

to the measurement of hedonic price indexes: hedonic dummy time variable indexes and hedonic imputation indexes. Both approaches not only correct price changes for changes in the quality of items purchased, but also allow the indexes to incorporate matched and unmatched models. However, they can yield quite different results and it is not immediately apparent which approach is preferable. The standard way price changes are measured by most national governments is through the use of the matched models method. In this method the details and prices of a representative selection of items are collected in a base reference period and their matched prices collected in successive periods so that the prices of 'like' are compared with 'like'. However, if there is a rapid turnover in available models, then the sample of product prices used to measure price changes becomes unrepresentative of the category as a whole as a result of both new unmatched models being introduced (but not included in the sample), and older unmatched models being retired (and thus dropping out of the sample). Hedonic indexes use matched and unmatched models and in doing so put an end to the matched models sample selection bias (Cole, *et al.* (1986), Silver and Heravi (2002) and Pakes (2003)).

Our interest in hedonic indexes arose from the need to reduce bias in the measurement of the U.S. consumer price index (CPI), which has been the subject of three major reports, the Stigler Committee (1961), Boskin Commission (1996) and the report by the Committee on National Statistics (2002): the Schultze panel. Each found the inability to properly remove the effect on price changes of changes in quality to be a major source of bias. Hedonic regressions were considered to be the most promising approach to control for such quality changes, though the Schultze panel cautioned for the need for further research on methodology.[211]

At first sight the two approaches to hedonic indexes appear quite similar and there is little to choose between them. Both rely on hedonic regression equations to remove the effects on price of quality changes. They can also incorporate a range of weighting systems, take specific functional forms for the aggregation (e.g. geometric and arithmetic) and can be formulated as chained or direct, fixed base comparisons. Yet they can give quite different results, even when using comparable weights, functional forms and the same periodic comparison. This is because they work on different principles. The dummy variable method constrains hedonic regression parameters to be the same over time while a hedonic imputation index paradoxically relies on parameter change as the essence of the measure. There has been some valuable research on the two approaches (see Berndt and Rap-

MacKinnon, 1993: 323). For a direct fixed base comparison the estimate reduces to a matched models one.

[211] "Hedonic techniques currently offer the most promising approach for explicitly adjusting observed prices to account for changing product quality. But our analysis suggests that there are still substantial unresolved econometric, data, and other measurement issues that need further attention." (Committee on National Statistics, 2002: 6).

paport (2001), Diewert (2002b), Silver and Heravi (2003), Pakes (2003) and Triplett (2004)), though no formal analysis, to the author's knowledge, of the factors governing the differences between the approaches. Berndt and Rappaport (2001) and Pakes (2003) have highlighted the fact that the two approaches can give different results and both advise the use of HI indexes when parameters are unstable, though our analysis casts more light on this proposal.

This paper first, in section I, examines the alternative formulations of the two methods and then, in section II develops an expression for their differences. Section III provides an empirical study for desktop PCs and section IV discusses the issue of choice between the approaches in light of the findings in sections II and II; section V concludes.

I. HEDONIC INDEXES

A hedonic regression equation of the prices of i=1,...,N models of a product, p_i, on their quality characteristics z_{ki} , where z_k =1,….,K price-determining characteristics, is given in a linear form by:

$$p_i = \gamma_0 + \sum_{k=1}^{K} \beta_k z_{ki} + \varepsilon_i$$

(1)

The β_k are estimates of the marginal valuations the data ascribes to each characteristic (Rosen (1974), Griliches (1988) and Triplett (1988), see also Diewert (2003) and Pakes (2003)). Statistical offices use hedonic regressions for CPI measurement when a model is no longer sold and a price adjustment for quality differences is needed in order that the price of the original model can be compared with that of a non-comparable replacement model. Silver and Heravi (2001) refer to this as 'patching'. However, patching can only make use of data outside of the matched sample when a model is missing. The replacements are on a one-to-one basis. The needs of quality adjustment, in dynamic markets, such as PCs, is to resample each month to cover a representative sample of all the new models in the market and *hedonic indexes* provide the required measures.

A. Hedonic imputation (HI) indexes

Hedonic imputation (hereafter—HI) indexes take a number of forms: first, as either equally-weighted or weighted indexes; second, depending on the functional form of the aggregator, say a geometric aggregator as against an arithmetic one; third, with regard to which period's characteristic set is held constant and finally, as direct comparisons between periods 0 and t, or as chained indexes, with individual links being calculated between periods 0 and 1, 1 and 2,….,t-1 and t, the results being combined by successive multiplication.

Consider the linear hedonic function $\hat{p}_i^0 = g^0\left(z_i^0\right)$ of (1) and a semi-logarithmic form of (1) $\hat{p}_i^0 = h^0\left(z_i^0\right)$ both estimated in period 0 with a vector of K quality characteristics $z_i^0 = z_{i1}^0,......,z_{iK}^0$ and N^0 observations and similarly for period t. Let quantities sold in period 0 be q_i^0 and relative sales value shares be $s_i^0 = p_i^0 q_i^0 / \sum_i p_i^0 q_i^0$ and again similarly for period t.

A hedonic Laspeyres index for *matched and unmatched period 0* models is given by:

$$P_{H-Las}=\frac{\sum_{i=1}^{N^0} g^t\left(z_i^0\right)q_i^0}{\sum_{i=1}^{N^0} g^0\left(z_i^0\right)q_i^0} \qquad (2)$$

and a hedonic Paasche index for *matched and unmatched period t* models by:

$$P_{H-Pas}=\frac{\sum_{i=1}^{N^t} g^t\left(z_i^t\right)q_i^t}{\sum_{i=1}^{N^t} g^0\left(z_i^t\right)q_i^t}\approx\frac{\sum_{i=1}^{N^t} p^t q_i^t}{\sum_{i=1}^{N^t} g^0\left(z_i^t\right)q_i^t} \qquad (3)$$

It is apparent from equations (2) and the first expression in (3) that a hedonic Laspeyres index holds characteristics constant in the *base* period while a hedonic Paasche index holds the characteristics constant in the *current* period. Thus the differences between the hedonic valuations in Laspeyres and Paasche are dictated by the extent to which the characteristics change over time; i.e. $\left(z_i^t - z_i^0\right)$. The further the z_i values are apart, say due to greater technological change, the less justifiable is the use of an individual estimate and the less faith there is in a compromise geometric mean of the two indexes—a Fisher index.[212] Since neither hedonic Laspeyres and nor Paasche indexes can be considered superior, we focus in the analytical work on superlative indexes—the Fisher, Törnqvist and Walsh formulas—which make *symmetric* use of information in *both* periods, are supported by economic theory and axiomatic considerations and are preferred target indexes in the forthcoming international CPI Manual (Diewert, 2004: chapters 15-18).[213]

For our consideration of matched *and* unmatched new and old models in the indexes we consider four exhaustive sets of models: $i\in S^t\left(0\cap t\right)$ matched models in period *t*; $i\in S^0\left(0\cap t\right)$ matched models in period 0; $i\in S^t\left(t\neg 0\right)$ unmatched new models present in period *t*, but not in period 0; $i\in S^0\left(0\neg t\right)$ unmatched old models present in period 0, but not in period *t*. Let the number of models in these respective sets be denoted by $N^t\left(0\cap t\right), N^0\left(0\cap t\right)$, $N^0\left(0\neg t\right)$ and $N^t\left(t\neg 0\right)$. We also denote the set of matched models with common characteristics $z_i^m = z_i^0 = z_i^t$ in both periods as $i\in S\left(0\cap t\right)$ enumerated over $i=1,\ldots,N^t\left(0\cap t\right)=N^0\left(0\cap t\right)=N^m$. A **generalized[214] hedonic Fisher index** (a geometric mean of (2) and (3)) is given by:

$$P_{H-Fisher}=\left[\left(\frac{\sum_{i\in S^0(0\cap t)}^{N^0(0\cap t)} g^t\left(z_i^m\right)q_i^0+\sum_{i\in S^0(0\neg t)}^{N^0(0\neg t)} g^t\left(z_i^0\right)q_i^0}{\sum_{i\in S^0(0\cap t)}^{N^0(0\cap t)} g^0\left(z_i^m\right)q_i^0+\sum_{i\in S^0(0\neg t)}^{N^0(0\neg t)} g^0\left(z_i^0\right)q_i^0}\right)\times\left(\frac{\sum_{i\in S^t(0\cap t)}^{N^t(0\cap t)} g^t\left(z_i^m\right)q_i^t+\sum_{i\in S^t(t\neg 0)}^{N^t(t\neg 0)} g^t\left(z_i^t\right)q_i^t}{\sum_{i\in S^t(0\cap t)}^{N^t(0\cap t)} g^0\left(z_i^m\right)q_i^t+\sum_{i\in S^t(t\neg 0)}^{N^t(t\neg 0)} g^0\left(z_i^t\right)q_i^t}\right)\right]^{\frac{1}{2}}$$

(4)

A generalized hedonic Törnqvist index is given by:

$$P_{H-Törnqvist}=\left[\frac{\prod_{i\in S^t(0\cap t)}^{N^m} h^t\left(z_i^m\right)}{\prod_{i\in S^0(0\cap t)}^{N^m} h^0\left(z_i^m\right)}\right]^{\tilde{s}_i}\times\left[\frac{\prod_{i\in S^0(0\neg t)}^{N^0(0\neg t)} h^t\left(z_i^0\right)}{\prod_{i\in S^0(0\neg t)}^{N^0(0\neg t)} h^0\left(z_i^0\right)}\right]^{\frac{s_i^0}{2}}\times\left[\frac{\prod_{i\in S^t(t\neg 0)}^{N^t(t\neg 0)} h^t\left(z_i^t\right)}{\prod_{i\in S^t(t\neg 0)}^{N^t(t\neg 0)} h^0\left(z_i^t\right)}\right]^{\frac{s_i^t}{2}} \qquad (5)$$

where $\tilde{s}_i = (s_i^0 + s_i^t)/2$.

A **generalized hedonic Walsh index** is given by:

---

[212] As an estimate of a COLI index the spread is irrelevant since the need is to include substitution effects and Fisher meets this need. However, Laspeyres and Paasche answer meaningful question and act as bounds on models of economic behaviour that different consumer might pursue. A Fisher estimate with less 'spread' is more satisfactory

[213] Pakes (2003) advocates a Laspeyres hedonic as an upper bound to a target compensating index but notes that a Paasche hedonic index may be more feasible in real time, again raising issues as to the extent of the difference.

[214] Diewert (2002b) referred to indexes that incorporated all unmatched and matched data as generalized forms.

$$P_{HBC-Walsh} = \frac{\sum_{i \in S(0 \cap t)}^{N^m} g^t\left(\sqrt{z_i^0 z_i^t}\right)\sqrt{q_i^0 q_i^t} + \sum_{i \in S^0(0 \neg t)}^{N^0(0 \neg t)} g^t\left(z_i^0\right)\sqrt{q_i^0} + \sum_{i \in S^t(t \neg 0)}^{N^t(t \neg 0)} g^t\left(z_i^t\right)\sqrt{q_i^t}}{\sum_{i \in S(0 \cap t)}^{N^m} g^0\left(\sqrt{z_i^0 z_i^t}\right)\sqrt{q_i^0 q_i^t} + \sum_{i \in S^0(0 \neg t)}^{N^0(0 \neg t)} g^0\left(z_i^0\right)\sqrt{q_i^0} + \sum_{i \in S^t(t \neg 0)}^{N^t(t \neg 0)} g^t\left(z_i^t\right)\sqrt{q_i^t}} \quad (6)$$

where $\sqrt{z_i^0 z_i^t} = z_i^m$ for the summation over N^m matched models.

B. Dummy time hedonic (DTH) indexes

Dummy time hedonic (hereafter—DTH) indexes are a second approach which, as with HI indexes, do not require a matched sample.[215] The formulation is similar to equation (1) except that a single regression is estimated on the data in the two time periods compared, $i \in S(0 \cup t)$. The equation also includes a dummy variable D^t which is equal to 1 in period *t*, and zero otherwise:

$$\ln p_i^{0,t} = \delta_0 + \delta_1 D^{0,t} + \sum_{k=1}^{K} \beta_k z_{ki}^{0,t} + \varepsilon_i^{0,t}$$

(7)

The exponent of the estimated coefficient δ_1^* is an estimate of the quality-adjusted price change between period 0 and period *t*. A weighted version of (7) would use a weighted least squares (WLS) estimator and weights $\tilde{s}_i^0 = (s_i^0 + s_i^t)/2$ for matched models and $s_i^0/2$ or $s_i^t/2$ for unmatched old and new models respectively (Diewert, 2002b).

The regression equation (7) constrains each of the β_k coefficients to be the same across the two months compared. In restricting the slopes to be the same, the (log of the) price change between periods 0 and *t* can be measured at any value of *z*, as illustrated by the difference between the dashed lines in Figure 1. For convenience it is evaluated at the origin as δ_1^*. Bear in mind the HI indexes outlined above estimate the differences between price surfaces with different slopes. As such the estimates have to be conditioned on particular values of *z*, which gives rise to the two estimates considered in (2) and (3): the base HI using z^0 and the current period HI using z^t, as shown in Figure 1. The very core of the DTH method is to constrain the coefficients to be the same, so there is no need to condition on particular values of *z*. The DTH estimates implicitly and usefully make symmetric use of base and current period data. As with hedonic imputation indexes DTH indexes can take fixed and chained base forms, though they can also take a fully constrained form whereby a single constrained regression is estimated for say January to December with dummy variables for each month, though this is impractical in real time since it requires data on future observations.

II. WHY HEDONIC IMPUTATION AND DUMMY TIME HEDONIC INDEXES DIFFER

A. Algebraic differences

(i) a reformulation of the hedonic indexes

There has been little analytical work undertaken on the factors governing differences between the two approaches. To compare the HI approach to the DTH approach we first need to reformulate the HI indexes. We note that the HI approach relies on two estimated hedonic equations, $h^t\left(z_i^t\right)$ and $h^0\left(z_i^0\right)$ for periods 0 and *t* respectively:

$$ln\ p_i^t = \gamma_0^t + \sum_{k=1}^{K} \beta_k^t z_{ki}^t + \varepsilon_i^t$$

(8)

$$\ln\ p_i^0 = \gamma_0^0 + \sum_{k=1}^{K} \beta_k^0 z_{ki}^0 + \varepsilon_i^0$$

(9)

We assume that the errors in each equation are similarly distributed, then phrase the two equations as a single hedonic regression equation with dummy time intercept and slope variables:

$$ln\ p_i^{0,t} = \gamma_0^0 + \gamma_1 D_0^{0,t} + \sum_{k=1}^{K} \beta_k^0 z_{ki}^{0,t} + \sum_{k=1}^{K} \beta_k D_{ki}^{0,t} + \varepsilon_i^{0,t}$$

(10)

where $D_0^{0,t} = 1$ if observations are in period *t* and 0 otherwise, $\gamma_1 = (\gamma_0^t - \gamma_0^0)$, $D_{ki}^{0,t} = z_{ki}^t$ if observations are in period *t* and 0 otherwise, and $\beta_k = (\beta_k^t - \beta_k^0)$. The γ_1^* is an estimate of the change in the *intercepts* of the two hedonic price equations and is thus a HI index evaluated at a particular value of $z_{ki}^{0,t}$, say denoted by $\tilde{z}_k^{0,t}$ which is zero at the intercept. A HI index evaluated at $\tilde{z}_k^{0,t} = 0$ has no economic meaning. We require a mean-value HI index evaluated at $\tilde{z}_k^{0,t} = \bar{z}_k^{0,t}$. The Walsh and Törnqvist indexes can be evaluated at mean values and, being superlative, should closely approximate each other and the Fisher index, all of which are recommended as target indexes for consumer price indexes (Diewert, 2004, chapters 15 and 17). The mean values that correspond to the generalized Törnqvist and Walsh HI indexes of (5) and (6) are respectively:[216]

$$\tilde{z}_k^{0,t} = \bar{z}_{k-Törn}^{0,t} = \prod_{i \in S(0 \cap t)}^{N^m} \left(z_i^m\right)^{\tilde{s}_i} \times \prod_{i \in S^0(0 \neg t)}^{N^0(0 \neg t)} \left(z_i^0\right)^{\frac{s_i^0}{2}} \times \prod_{i \in S^t(t \neg 0)}^{N^t(t \neg 0)} \left(z_i^t\right)^{\frac{s_i^t}{2}}$$

(11)

[215] See De Haan (2003) for a variant that uses matched data when available and the time dummy only for unmatched data.

[216] Mean value weights that apply to price *changes* differ from those that apply to prices in the Walsh index. The former are $\sqrt{q_i^0 q_i^t p_i^0} \Big/ \Sigma i \sqrt{q_i^0 q_i^t p_i^0}$ for matched models, as opposed to $\sqrt{q_i^0 q_i^t} \Big/ \Sigma i \sqrt{q_i^0 q_i^t}$, but it is the latter that concern us.

$$\tilde{z}_k^{0,t} = \bar{z}_{k-Walsh}^{0,t} = \frac{\sum_{i \in S(0 \cap t)}^{N^m} \left(z_i^m\right)\sqrt{q_i^0 q_i^t} + \sum_{i \in S^0(0 \neg t)}^{N^0(0 \neg t)} \left(z_i^0\right) q_i^0 + \sum_{i \in S^t(t \neg 0)}^{N^t(t \neg 0)} \left(z_i^t\right) q_i^t}{\sum_{i \in S(0 \cap t)}^{N^m} \sqrt{q_i^0 q_i^t} + \sum_{i \in S^0(0 \neg t)}^{N^0(0 \neg t)} q_i^0 + \sum_{i \in S^t(t \neg 0)}^{N^t(t \neg 0)} q_i^t}$$

(12)

We thus do not evaluate γ_1 at $\tilde{z}_k^{0,t} = 0$, but for Törnqvist and Walsh HI indexes at $\tilde{z}_k^{0,t} = \bar{z}_{k-Törn}^{0,t}$ and $\tilde{z}_k^{0,t} = \bar{z}_{k-Walsh}^{0,t}$ respectively. The required **Törnqvist HI index is thus estimated** (hereafter all references to indexes are after taking the exponents) as:

$$\gamma_1^* + \sum_{k=1}^{K} \beta_k^* \bar{z}_{k-Törn}^{0,t}$$ where β_k^* is a WLS[217] estimate of $(\beta_k^t - \beta_k^0)$

(13)

Figure 1 illustrates for a single (k=1) z variable how the estimate of the difference between the two hedonic equations at the intercept, γ_1^*, has an adjustment to allow it to be evaluated at the Törnqvist mean from (11), $\bar{z}_{k-Törn}^{0,t}$. Note that we do not estimate the equations using OLS, but an estimator whose weights correspond to a Törnqvist index, i.e., a WLS where the weights are those outlined after equation (7). We focus on the case of the Törnqvist HI index with similar principles applying to the Walsh HI index.

Consider now the DTH index in (7) which constrains $\beta_k = (\beta_k^t - \beta_k^0) = 0$ in (10) and thus $\sum_{k=1}^{K} \beta_k D_k^{0,t}$ to be zero. The DTH index in (7) corresponds to a **Törnqvist (DTH) index** if estimated using WLS where the weights are those outlined after equation (7) above. A natural question is how does the estimated DTH index δ_1^* in (7), which is invariant to values of $\boldsymbol{z}_{ki}^{0,t}$, differ from the HI index evaluated *at the means* of $\boldsymbol{z}_{ki}^{0,t}$ in (13)?

(ii) How does a Törnqvist HI index differ from a Törnqvist DTH index?

This difference is first considered by comparing $\left(\delta_1^* - \gamma_1^*\right)$, the difference *in the intercepts,* where $\tilde{z}_k^{0,t} = 0$, between the constrained and unconstrained regression equations (7) and (10) respectively. The difference is 'omitted variable bias' due to the omission of $\sum_{k=1}^{K} \beta_k D_k^{0,t}$ in (7). Consider the case of a single k=1 characteristic, the principles being readily extended to more variables. Consider further the regression, using data in periods 0 and t, of the omitted variable—the slope dummy variable, $D_{1i}^{0,t} = z_{1i}^t$ if period t and 0 otherwise, in (10)—on the remaining right-hand side variables, the intercept dummy $D_0^{0,t}$, the $z_{1i}^{0,t}$ characteristics and a disturbance term

[217] The weights used to correspond to a generalized Törnqvist index were outlined following equation (7) .

side variables, the intercept dummy $D_0^{0,t}$, the $z_{1i}^{0,t}$ characteristics and a disturbance term $\omega_i^{0,t}$ with desirable properties:

$$D_{1i}^{0,t} = \lambda_0^0 + \lambda_1 D_0^{0,t} + \lambda_2^0 z_{1i}^{0,t} + \omega_i^{0,t}$$

(14)

Omitted variable bias is the product of the coefficient on the omitted variable, β_1 (for *k*=1 in (10)), and the coefficient λ_1^* which is estimated from the above regression (14) on $D_{1i}^{0,t}$ i.e. the multicollinearity between the missing and included variables (Davidson and McKinnon, 1993). Thus the difference (before taking exponents) DTH *minus* HI, *at the intercept* is $\left(\delta_1^* - \gamma_1^*\right) = \beta_1^* \times \lambda_1^*$ and the (log of the) DTH index is thus given by $\delta_1^* = \beta_1^* \times \lambda_1^* + \gamma_1^*$. The (log of the) Törnqvist HI index at $\bar{z}_{1-Törn}^{0,t}$ from (11) for one variable is estimated as:

$$\gamma_1^* + \beta_1^* \bar{z}_{1-Törn}^{0,t}$$

(15)

Thus the (log of the) Törnqvist DTH index *minus* the (log of the) Törnqvist HI index at $\bar{z}_{1-\text{Törn}}^{0,t}$ is:

$$\beta_1^* \times \lambda_1^* + \gamma_1^* - \left(\gamma_1^* + \beta_1^* \bar{z}_{1-\text{Törn}}^{0,t}\right) = \beta_1^* \left(\lambda_1^* - \bar{z}_{1-\text{Törn}}^{0,t}\right) = \left(\beta_1^{*t} - \beta_1^{*0}\right)\left(\lambda_1^* - \bar{z}_{1-\text{Törn}}^{0,t}\right)$$

(16)

Alternatively the ratio of the DTH and HI indexes *at the intercept* is $exp\left(\delta_1^*\right) / exp\left(\gamma_1^*\right) = exp\left(\beta_1^* \times \lambda_1^*\right)$ and the DTH index is thus given by $exp\left(\delta_1^*\right) = exp\left(\beta_1^* \times \lambda_1^*\right) \times exp\left(\gamma_1^*\right)$. The Törnqvist HI index at $\bar{z}_{1-Törn}^{0,t}$ from (11) for one variable is estimated as: $exp\left(\gamma_1^* + \beta_1^* \bar{z}_{1-\text{Törn}}^{0,t}\right)$. Thus the ratio of the Törnqvist DTH index to the Törnqvist HI index at $\bar{z}_{1-\text{Törn}}^{0,t}$ is:

$$\left(\frac{\text{Törnqvist DTH index}}{\text{Törnqvist HI index}}\right)_{:\bar{z}_{1-\text{Törn}}^{0,t}} = exp\left[\left(\beta_1^{*t} - \beta_1^{*0}\right)\left(\lambda_1^* - \bar{z}_{1-\text{Törn}}^{0,t}\right)\right]$$

(17)

where β_1^{*t} and β_1^{*0} are WLS estimates.

If *either of the two terms* making up the product on the right–hand-side is close to zero then there will be little difference between the indexes. *Neither parameter instability nor a change in the mean characteristic is sufficient in itself to lead to a difference between the formulas*. The $\beta_1^* = \left(\beta_k^{*t} - \beta_k^{*0}\right)$ from (10) is the estimated marginal valuation of the characteristic between periods 0 and *t*, which can be positive or negative, but may be more generally thought to be negative to represent diminishing marginal utility/cost of the characteristic.

We now consider the WLS estimated λ_1^*. Bear in mind that the left-hand-side of the regression in equation (14), $D_{1i}^{0,t}$, is 0 in period 0 and z_{1i}^t in period *t* and that D^t on the right-hand-side is 1 in period *t* and zero in period 0. If we assume quality characteristics are positive, λ_1 will always be positive as the change from 0 in period 0 to their values in period *t*. Consider the weighted (Törnqvist) means $\bar{z}_{1-Torn}^0$ for $i \in S^0(1 \cap 2)$ and $i \in S^0(0 \neg t)$ (matched period 0 and unmatched old period 0) and $\bar{z}_{1-Torn}^t$ for $i \in S^t(1 \cap 2)$ and $i \in S^t(t \neg 0)$ (matched period *t* and unmatched new models in period *t)*.

If we assume for simplicity that $\bar{z}_{1-Törn}^0 = \bar{z}_{1-Törn}^t$, then $\lambda_1^* \approx \bar{z}_{1-Torn}^t$ since it is an estimate of the change in $D_{1i}^{0,t}$ in (14) arising from changing from period 0, where $D_{1i}^{0,t} = 0$, to period *t* where it is $\bar{z}_{1-Torn}^t$ and has an expectation of $\bar{z}_{1-Törn}^t$. The λ_1^* estimate is conditioned in (14) on $\bar{z}_{1-\text{Törn}}^{0,t}$, the change from $\bar{z}_{1-Törn}^0$ to $\bar{z}_{1-Törn}^t$, but since we assume these to have not changed our estimate of $\lambda_1^* \approx \bar{z}_{1-Törn}^t$ holds true.[218] Thus the second part of the difference expression in (14), $\left(\lambda_1^* - \bar{z}_{1-Törn}^{0,t}\right)$, is simply $\left(\bar{z}_{1-Törn}^t - \bar{z}_{1-Törn}^{0,t}\right)$ which, given our assumption of $\bar{z}_{1-Törn}^0 = \bar{z}_{1-Törn}^t$, is equal to 0. Thus from the right-hand side of (17), for samples with negligible change in the mean values of the characteristics, the DTH and HI will be similar irrespective of any parameter instability. Diewert (2002b) and Aizcorbe (2003) have shown that the DTH and HI indexes will be the same for matched models and this analysis gives support to their finding. However, we find first, that it is not matching *per se* that dictates the relationship; for unmatched models all that is required is that $\bar{z}_{1-\text{Törn}}^t = \bar{z}_{1-\text{Törn}}^{0,t}$ which may occur without matching—it simply requires the means of the characteristics not to change. Second, that even when the means change the two approaches will be equal if $\beta_k^{*t} - \beta_k^{*0} = 0$, i.e. there is parameter stability. Finally, it follows that if either of the two right-hand side expressions in (17) are large, the differences between the indexes will be compounded.

But what if $\bar{z}_{1-Törn}^0 \neq \bar{z}_{1-Törn}^t$**?** The estimated coefficient λ_1^* from (14) regression is given by:[219]

$$\lambda_1^* = \frac{\bar{z}_1^t (N - N^t) \sigma_z^2 - N \, cov(D_{1i}^{0,t}, z_{1i}^{0,t})(\bar{z}_1^t - \bar{z}_1^{0,t})}{(N - N^t)\sigma_z^2 - (\bar{z}_1^t - \bar{z}_1^{0,t})^2 N^t}$$

(18)

---

[218] We are not requiring $\bar{z}_{1-Törn}^0 = \bar{z}_{1-Törn}^t$; unconditioned estimates might be evaluated by examining omitted variable bias in (12).

[219] The estimated coefficient on x_1 of a regression of y on x_1 and x_2 is given by:

$$\frac{\Sigma yx_1 \Sigma x_2^2 - \Sigma yx_2 \Sigma x_1 x_2}{\Sigma x_1^2 \Sigma x_2^2 - \left(\Sigma x_1 x_2\right)^2}$$

for an unweighted regression where N^t and N are the respective number of observations in period *t* and both periods 0 and *t*, σ_z^2 is the variance of *z* and $cov(D_{1i}^{0,t}, z_{1i}^{0,t}) = \sum (z^t)^2 - N^t \bar{z}^t \bar{z}$ is the covariance of $D_{1i}^{0,t}$ and $z_{1i}^{0,t}$ from (14). Readers are reminded that from (17):

$$\left(\frac{\text{Törnqvist DTH index}}{\text{Törnqvist HI index}} \right)_{:\bar{z}_{1-\text{Törn}}^{0,t}} = exp\left[(\beta_1^{*t} - \beta_1^{*0})\left(\lambda_1^* - \bar{z}_{1-\text{Törn}}^{0,t}\right) \right]$$

(19)

First, as noted, if there is *either* negligible parameter instability *or* a negligible change in the mean of the characteristic, then there will be little difference between the formulas. However, as parameter instability increases *and* the change in the mean characteristic increases, the multiplicative effect on the difference between the indexes is compounded. The likely direction and magnitude of any difference is not immediately obvious. Assume diminishing marginal valuations of characteristics, so that $(\beta_1^{*t} - \beta_1^{*0}) \langle 0$. Second, even assuming a positive technological advance, $(\bar{z}_1^t - \bar{z}_1^{0,t}) \rangle 0$, and given $(N - N^t)$, $\text{cov}(D_{1i}^{0,t}, z_{1i}^{0,t})$, $\bar{z}_1^t$ and σ_z^2 are positive, it remains difficult to establish from (18) the effect on λ_1^* of changes in its constituent parts. However, third, as N^t becomes an increasing share of N, and at the limit if N^t takes up all of N (i.e. $(N - N^t) = 0$), then $\text{cov}(D_{1i}^{0,t}, z_{1i}^{0,t}) \to \sigma_z^2$ and more importantly, $\bar{z}_1^t \to \bar{z}_1^{0,t}$ and the difference between the formulas is dictated by $(\beta_1^{*t} - \beta_1^{*0})\left(-\bar{z}_1^{0,t}\right)$; i.e. a DTH index will exceed a HI index. Note that (18) is based on an OLS estimator and for a WLS Törnqvist estimator similar principles apply, though the determining factor for a DTH index to exceed a HI index is for the weights of new models to be increasing, a much more reasonable scenario. Thus other things being equal, the nature and extent of any differences between the two indexes will depend on (i) changes in the mean quality of models $(\bar{z}_1^t - \bar{z}_1^{0,t})$, (ii) the relative number of models in each period, $(N - N^t)$, (iii) the dispersion in *z*, σ_z^2, (iv) the absolute mean of the characteristics in period *t*, $\bar{z}_1^t$, (v) the $\text{cov}(D_{1i}^{0,t}, z_{1i}^{0,t})$ which as $(N - N^t) \to 0$, i.e. N^t takes up all of N, then $\text{cov}(D_{1i}^{0,t}, z_{1i}^{0,t}) \to \sigma_z^2$ and $\bar{z}_1^t = \bar{z}_1$ and $\lambda_1^* = 0$, and (vi) parameter instability, $(\beta_1^{*t} - \beta_1^{*0})$.

B. Treatment of unmatched observations

Diewert (2002b) and Aizcorbe (2003) show that while the DTH and HI indexes will be the same for matched models, they differ in their treatment of unmatched data. Consider hedonic functions $\boldsymbol{h}_i^t\left(z_i^t\right)$ and $\boldsymbol{h}_i^0\left(z_i^0\right)$ for periods *t* and 0 respectively as in (8) and (9) and a (constrained) time dummy regression equation (7). Consider an unmatched observation only available in period *t*. A base period HI index such as (2) would exclude it, while a current period HI index such as (3) would include it and a geometric mean of the two (4) would give it half the weight in the calculation of that of a matched observation. A Törnqvist hedonic index (5) would also give an unmatched model half the weight of a matched one. For a DTH index such as (7) an unmatched period *t* model would appear only once in period *t* in the estimation of constrained parameters, as opposed to twice for matched data. We would therefore expect superlative HI indexes, such as (4), to be closer to DTH indexes than their constituent elements, (2) and (3) because they make symmetric use of the data.

C. Observations with undue influence

Silver (2002) has shown that while HI indexes such as (2)-(6) explicitly incorporate weights, they are only implicitly incorporated in the OLS or WLS estimator used for DTH indexes in a manner that may not be fully representative of the weights used. Adverse influence and leverage effects are shown to be generated by observations with unusual characteristics and above average residuals.

D. Chaining

Chained base HI indexes are preferred to fixed base ones, especially when matched samples degrade rapidly and their use reduces the spread between Laspeyres and Paasche. However, caution is advised in the use of chained monthly series when prices may oscillate around a trend (i.e. 'bounce') and as a result, chained indexes can 'drift' (Forsyth and Fowler, 1981 and Szulc, 1983).

III. EMPIRICAL STUDY: DESKTOP PCs

The empirical study is of the measurement of quality-adjusted monthly prices of UK desktop PCs in 1998. The data were monthly scanner data from the bar-code readers of PC retailers. There were 7,387 observations (models in a month sold in a specialized or non-specialized PC store-type) representing a sales volume of 1.5 million models worth £1.57 billion pounds over the year. Table 1 shows that in the January to February price comparison there were 584 matched models available in both months for the price comparison. However, for the January to December price comparison only 161 matched models were available with 509 unmatched 'old' models (available in January, but unmatched in December) and 436 unmatched 'new' models (available in December but unavailable in January for matching).[220]

The calculation of hedonic indices first require the estimation of hedonic regressions. To simplify the illustration only one variable was included, the speed in MHz. The regressions were for each month for the HI indexes and over January and the current month, including a dummy variable for the latter, for the DTH indexes. The estimated coefficients for speed in the hedonic regressions were statistically significant coefficients with the expected positive signs. [221]

The selectivity bias inherent in a *matched models* Törnqvist index is shown in Table 1 by the fall of 24% which understates the fall of around 50% for the two Törnqvist *hedonic* indexes which use all

---

[220] Bear in mind that the indexes estimated in this paper are weighted by shares in sales values and that the fall-off in the coverage of the matched sample by sales is even more dramatic: for the January to December comparison matched models made up only 71% of the January sales value and a mere 12% of the December sales value.

[221] The F-statistics for the null hypothesis of coefficients being equal to zero averaged 34.2 for HI indexes and 53.4 for DTH indexes, consistently rejecting the null at a 0.01% level and lower. The explanatory power of the estimated equations were naturally low for this specification with a single explanatory variable, especially since they did not include dummy variables on brand. Details of estimates from a fully specified model are available from the authors.

of the data. The particular concern of this paper is that the HI and DTH Törnqvist indexes give different results: falls of 55% and 50% respectively—a difference of nearly 10%—with no immediate explanation as to which is better. Having shown the need for hedonic indices we now turn to an explanation in Table 2 of the differences between the results from the two methods.

Table 2 column (1) shows the ratio of the two estimates, finding the DTH estimate to consistently fall at a slower rate. The early months have very little quality-adjusted price change compared with January and the difference between the formulas are less amenable to analysis. However, from later months it is clear that while the parameter instability $\left(\beta_1^{*t} - \beta_1^{*0}\right)$ in column (2) is an important driver of the differences between the hedonic indexes, other factors are at work. For example, in September and December the differences between the parameters were about the same, yet the differences between the HI and DTH indexes in column (1) are higher in December than September. Furthermore, in October and November the parameter instability falls compared with September, yet the differences between the indexes increase. The other factor at work and affecting the difference, as shown in equation (17), is of course $\left(\lambda_1^* - \bar{z}_{1-Törn}^{0,t}\right)$ which in column (3) of Table 2 is seen to increase (in absolute values) in October, November and December. This increase explains that part of the differences in hedonic indexes unexplained by the parameter instability. The product of the two terms in column (3) shows an overall increase in October, November and December and, indeed, the exponent of column (3) equals column (1) by equation (17).

A driver of the substantial fall in $\left(\lambda_1^* - \bar{z}_{1-Törn}^{0,t}\right)$ can be seem from columns (5) and (6) to be relatively large increase in $\bar{z}_{1-Törn}^{0,t}$, the average speed of PCs over the two periods compared, an increase from 208 MHz for the January to February comparison to 273 MHz. for the January to December comparison. Bear in mind that the indexes were estimated using Törnqvist weights as outlined following equation (7). The overall sales value weight given to December models compared with January models increased by $\left[1-\left(0.70736-2\right)=0.29\right]$ 29% and the resulting fall in the statistic in column 9 of Table 2 is at least one of the reasons why the DTH index exceeded the HI index. Equation (18) and columns (7) to (10) of Table 2 show how λ_1^* can be decomposed, though the complexity of equation (18) makes it difficult to explain the factors that dictate its change, unless $\bar{z}_1^t = \bar{z}_1^{0,t}$ which is clearly not the case.

IV. CHOICE BETWEEN HEDONIC INDEXES AND TIME DUMMY HEDONIC INDEXES

The question of choice between these approaches has mainly been considered[222] from a perspective of concern over parameter instability. Berndt and Rappaport (2001) found, for example, from 1987

[222] Diewert (2002b) points out that the main advantage of hedonic imputed indexes is that it they are more flexible; i.e., changes in tastes between periods can readily be accommodated. HI indexes are argued to have a *disadvantage* that *two* distinct estimates will be generated and it is somewhat arbitrary how these two estimates are to be averaged to form a

to 1999 for desktop PCs, the null hypothesis of adjacent-year equality to be rejected in all but one case. For mobile PCs the null hypothesis of parameter stability was rejected in eight of the 12 adjacent-year comparisons. Berndt and Rappaport (2001) preferred the use of HI indexes if there was evidence of parameter instability. Pakes (2003) using quarterly data for hedonic regressions for desktop PCs over the period 1995 to 1999 rejected just about any hypothesis on the constancy of the coefficients. He also advocated HI indexes on the grounds that "....since hedonic coefficients vary across periods it [the DTH index approach] has no theoretical justification." Pakes (2003: 1593). However, equation (19) showed how the ratio of DTH and HI indexes was not solely dependent on parameter instability. It depended on the exponent of the product of two components: the change over time in the (WLS estimated) hedonic coefficients *and* the difference in (statistics that relate to) the (weighted) mean values of the characteristic. Even if parameters were unstable, the difference between the indexes may still be small due to a minimal change in the other component.

Consider two forms of 'spread'. The first arises because the basis of a DTH index is the constraining of parameters over time to be the same. Constraining the estimated coefficients to either β_1^{*t} or β_1^{*0} may give quite different results and this difference is the spread. A DTH index constrains the parameters to be the same, an average of the two. There is a sense in which we have less confidence in an index based on constraining similar parameters than one based on constraining two disparate parameters.

But there is a second type of spread and this relates to HI indexes. This is the difference between base and current period hedonic indexes, between (2) and (3), and arises from the use of a constant z_i^t as against z_i^0. The difference or spread between the indexes is dictated by the choice between which period's characteristics are held constant and there is a sense in which we have more confidence in an index based on an average of z_i^0 and z_i^t that has relatively small spread, than one based on an average of z_i^0 and z_i^t in which they were very different. There is thus a case to avoid HI indexes and use DTH indexes when characteristics change. If neither $(\beta_1^{*t} - \beta_1^{*0})$ nor $\left(\lambda_1^* - \bar{z}_{1-\text{Törn}}^{0,t}\right)$ is particularly large[223] relative to the other, then a symmetric average, say geometric mean, of the two indexes is to be preferred. If both differences are significant, but differ, a more appropriate estimate might be a weighted mean such as:

$$\frac{exp\left(\beta_1^{*t} - \beta_1^{*0}\right) \times HI + exp\left(\lambda_1^* - \bar{z}_{1-\text{Törn}}^{0,t}\right) \times DTH}{exp\left(\beta_1^{*t} - \beta_1^{*0}\right) + exp\left(\lambda_1^* - \bar{z}_{1-\text{Törn}}^{0,t}\right)}$$

(20)

---

single estimate of price change. Diewert (2002b) further identifies the main advantages of the DTH method are that it conserves degrees of freedom and is less subject to multicollinearity problems.

V. CONCLUSIONS

It is recognised that extensive product differentiation with a high model turnover is an increasing feature of product markets (Triplett, 1999). The motivation of this paper lay in the failure of the matched models method to adequately deal with price measurement in this context and the need for hedonic indexes as the most promising alternative (Schultze and Mackie, 2002). The paper first developed in section I *generalized superlative hedonic indexes*, i.e. index number formulas which are *generalized* to deal with matched and unmatched models, use *hedonic* regressions to control for quality changes and are *superlative* in that they have a good foundations in economic theory and axiomatic considerations (Diewert, 2004). The paper second, considered the two main approaches to hedonic index numbers, HI, and DTH indexes. Their commonalities in terms of functional form of aggregator, weighting and periodicity of comparison, makes it difficult to justify one approach against the other, yet the two approaches can yield quite different results. This is of concern if they are to be the principle tools for dealing with quality adjustment in price measurement for product markets with high model turnover. In section II the paper provided a formal exposition of the factors underlying the difference between the two approaches. It was shown that differences between the two approaches may arise from both parameter instability and changes in the characteristics and such differences are compounded when both occur. It further showed that similarities between the two approaches resulted if there was little difference in either component. The empirical study of desktop PCs showed the superiority of hedonic indexes over matched model indexes, that DTH and HI indexes can differ, and that parameter instability need not be the main factor dictating such differences.

Consideration of the issue of choice between the two approaches was based in section IV on minimising two concepts of spread. The analysis led to the advice that (i) Either the DTH or HI index approach is acceptable if either the parameters are relatively stable or the characteristics do not change over time, otherwise (ii) DTH indexes be avoided and HI indexes used when there is evidence of parameter instability, (iii) HI indexes be avoided and DTH indexes used when there is evidence of changes in quality characteristics of the form given in the last term of (19), and (iv) that when both characteristics and parameters change substantially, an average, such as (20), of the two approaches be used.

[223] Note that statistical significance tests should not dictate whether a difference is large or small, it is also the magnitude of the differences that matters.

References

Aizcorbe, A. (2003) The Stability of Dummy Variable Price Measures Obtained from Hedonic Regressions, *Mimeo*, Federal reserve Board, Washington, D.C.

Berndt, E.R. and Rappaport N.J. (2001). Price and Quality of Desktop and Mobile Personal Computers: A Quarter-Century Historical Overview, *American Economic Review*, 91, 2, 268-273.

Boskin MS (Chair) Advisory Commission to Study the Consumer Price Index (1996), Towards a More Accurate Measure of the Cost of Living, *Interim Report to the Senate Finance Committee*, Washington DC.

Cole, R., Chen, Y.C., Barquin-Stolleman, J.A., Dulberger, E., Helvacian, N. and Hodge, J.H., (1986), Quality-Adjusted Price Indexes for Computer Processors and Selected Peripheral Equipment, *Survey of Current Businesses*, 66, 1, January 41-50.

Committee on National Statistics (2002) At What Price? Conceptualising and Measuring Cost-of-Living and price indexes, Panel on Conceptual, Measurement and Other Statistical Issues in Developing Cost-of-Living Indexes, Charles Schultze and Chris Mackie (eds.), Committee on National Statistics, Washington D.C., National Academy Press.

Davidson, J.R. and Mackinnon, J.G. (1993) *Estimation and Inference in Econometrics*, Oxford University Press.

Diewert, W.E. (2002a), Weighted Country Product Dummy Variable Regressions and Index Number Formulae, *University of British Columbia Department of Economics Discussion Paper Series*, 02-15, September, Vancouver, B.C., Canada.

Diewert W.E. (2002b). Hedonic Regressions: A Review of Some Unresolved Issues, *Mimeo*, Department of Economics, University of British Columbia.

Diewert, W.E. (2003), Hedonic Regressions: A Consumer Theory Approach. In Mathew Shapiro and Rob Feenstra (eds.), *Scanner Data and Price Indexes*, National Bureau of Economic Research, Studies in Income and Wealth, vol.61 Chicago: University of Chicago Press, 317-348.

Diewert, W.E. (2004), Chapters 15 to 18 in: The Manual on Consumer Price Indices: Theory and Practice, International Labour Office, Geneva: ILO.

Forsyth, FG, and Fowler, RF (1981), The Theory and Practice of Chain Price Index Numbers, *Journal of the Royal Statistical Society*, A, 144, 2, 224-47.

Griliches, Zvi [1988], "Postscript on Hedonics," in Zvi Griliches, *Technology, Education, and Productivity*, New York: Basil Blackwell, Inc., 119-122.

Haan, J. de (2003), Time Dummy Approaches to Hedonic Price Measurement, *Paper presented at the Seventh Meeting of the International Working Group on Price Indices*, 27-29 May 2003, INSEE, Paris.

Pakes A. (2003), A Reconsideration of Hedonic Price Indexes with an Application to PCs, *The American Economic Review*, 93, 5, December, 1576-93.

Rosen, S. (1974), Hedonic Prices and Implicit Markets: Product Differentiation and Pure Competition, *Journal of Political Economy*, 82, 34-49.

Schultze,C. and Mackie C. (2002) (editors), see Committee on National Statistics (2002) *op. cit.*

Silver, M. and Heravi, S. (2001), Scanner Data and the Measurement of Inflation, *The Economic Journal*, 11 June, 384-405.

Silver, M and Heravi, S. (2002), Why the CPI Matched Models Method May Fail Us: Results from an Hedonic and Matched Experiment Using Scanner Data, European *Central Bank (ECB) Working Paper* No. 144, Frankfurt: ECB.

Silver, M. (2002), The Use of Weights in Hedonic Regressions: the Measurement of Quality-Adjusted Price Changes. *Mimeo*, Cardiff Business School, Cardiff University.

Silver, M. and Heravi, S. (2003), The Measurement of Quality-Adjusted Price Changes. In Mathew Shapiro and Rob Feenstra (eds.), *Scanner Data and Price Indexes*, National Bureau of Economic Research, Studies in Income and Wealth, vol. 61, Chicago: University of Chicago Press, 277-317.

Stigler, G. (1961) The Price Statistics of the Federal Government, *Report* to the Office of Statistical Standards, Bureau of the Budget, New York, NBER.

Szulc, BJ (1983) Linking Price Index Numbers, in *Price Level Measurement*, 537-566, Ottawa, Statistics Canada.

Triplett, J.E. (1988), Hedonic Functions and Hedonic Indexes, in *The New Palgraves Dictionary of Economics*, pp. 630-634.

Triplett, J.E. (1999) The Solow Productivity Paradox: What do Computers do to Productivity?, *Canadian Journal of Economics*, 32, 2, April , 309-334.

Triplett, J.E. (2004 - forthcoming) Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes. Directorate for Science, Technology and Industry, Paris: OECD

Jan de Haan: **Virhe. Viitteen lähdettä ei löytynyt.**

Abstract: This paper compares a Törnqvist price index in which the 'missing prices' are imputed using hedonic regression with the time dummy hedonic index. The aim is to show that the time dummy index can be interpreted as a special case of the imputation Törnqvist index when the regression weights are properly chosen. It is argued that the set of weights proposed in the recent literature overstates the impact of new and disappearing items. This could be particularly relevant for high-turnover goods like PCs.

Keywords: consumer price index; hedonic imputation; time dummy index.

Introduction

Silver and Heravi (2004) address the difference between hedonic imputation indexes and time dummy variable hedonic indexes, which are the two main approaches in the academic literature to estimating hedonic price indexes. The larger part of their paper relates to the imputation Törnqvist index and the time dummy index that uses the regression weights proposed by Diewert (2003). They analyse the factors driving the difference between both approaches and show that differences may arise from parameter instability and changes in the average characteristics and such differences are compounded when both occur. Their idea of looking at the time dummy variable model – which constrains the parameters to be the same across the two periods – as one that suffers from 'omitted variable bias' is particularly interesting.

Silver and Heravi define imputation price indexes as indexes in which all prices are estimated using an hedonic model. That is, in addition to the imputation of 'missing prices' of new or disappearing items, observed prices are replaced by their predicted values. But throwing away observed prices is not an attractive idea, certainly not for statistical agencies. This paper therefore focuses on a (partial) imputation Törnqvist index in which hedonic imputation is restricted to 'missing prices'. It is shown that the time dummy approach can be interpreted as a special case of the partial hedonic imputation Törnqvist index when the regression weights are properly chosen.

The paper is organized as follows. Section 2 formally defines the partial imputation Törnqvist index. Section 3 shows that imputing 'missing prices' is essentially what a proper time dummy index does. Section 4 argues that the weights Diewert (2003) has suggested for the unmatched items in the estimation of the time dummy index are twice as large as they should be, at least if the partial imputation Törnqvist index serves as the preferred target index. Section 5 addresses some econometric issues. Section 6 provides empirical evidence using scanner data on personal computers for the Netherlands. Section 7 concludes.

The partial hedonic imputation Törnqvist index

I start by introducing some notation. Let p_i^0 and p_i^1 denote the price of item *i* in the base period 0 and the current or comparison period 1, respectively, and let s_i^0 and s_i^1 denote the corresponding expenditure shares or relative sales values. Further, let S^0 and S^1 be the sets of models available in the respective periods. $S_M = S^0 \cap S^1$ denotes the set of matched items; S_D^0 is the set of disappearing items (those sold in period 0 but no longer in period 1) and S_N^1 the set of new items (those sold in period 1 but not in period 0). First I will define what I call the *full hedonic imputation* (FHI) geometric Laspeyres (GL) and geometric Paasche (GP) indexes. They can be written as

$$P_{FHIGL} = \prod_{i \in I^0} \left(\frac{\tilde{p}_i^1}{\tilde{p}_i^0} \right)^{s_i^0} = \prod_{i \in S_M} \left(\frac{\tilde{p}_i^1}{\tilde{p}_i^0} \right)^{s_i^0} \prod_{i \in S_D^0} \left(\frac{\tilde{p}_i^1}{\tilde{p}_i^0} \right)^{s_i^0}, \qquad (1)$$

and

$$P_{FHIGP} = \prod_{i \in S^1} \left(\frac{\tilde{p}_i^1}{\tilde{p}_i^0} \right)^{s_i^1} = \prod_{i \in S_M} \left(\frac{\tilde{p}_i^1}{\tilde{p}_i^0} \right)^{s_i^1} \prod_{i \in S_N^1} \left(\frac{\tilde{p}_i^1}{\tilde{p}_i^0} \right)^{s_i^1}, \qquad (2)$$

where $\tilde{p}_i^t$ denotes the predicted price of *i* in period *t* (*t*= 0,1), estimated with an hedonic model.[224] Note that the prices in the expenditure shares are not replaced by their predicted values. Expressions (1) and (2) are *generalized* indexes in the sense that they incorporate all unmatched and matched items (see Diewert, 2003). There are two (related) problems with using those indexes. First, only in special circumstances will they coincide with the indexes based on observed prices when there are no unmatched items, depending on the specification of the model and the regression weights used. Second, statistical agencies may be reluctant to replace observed prices by predicted values, and to my opinion they are right. I would prefer an approach that only imputes $\tilde{p}_i^1$ for $i \in S_D^0$ and $\tilde{p}_i^0$ for $i \in S_N^1$, leaving all observed prices unchanged, because this minimizes the impact of econometric modelling. This is what statistical agencies usually mean by imputation anyway.[225] The *partial hedonic imputation* (PHI) geometric Laspeyres index and geometric Paasche index are given by

$$P_{PHIGL} = \prod_{i \in S_M} \left(\frac{p_i^1}{p_i^0} \right)^{s_i^0} \prod_{i \in S_D^0} \left(\frac{\tilde{p}_i^1}{p_i^0} \right)^{s_i^0}, \qquad (3)$$

and

$$P_{PHIGP} = \prod_{i \in S_M} \left(\frac{p_i^1}{p_i^0} \right)^{s_i^1} \prod_{i \in S_N^1} \left(\frac{p_i^1}{\tilde{p}_i^0} \right)^{s_i^1}. \qquad (4)$$

Taking the geometric average of (3) and (4) leads to the partial hedonic imputation Törnqvist index:

$$P_{PHIT} = \prod_{i \in S_M} \left(\frac{p_i^1}{p_i^0} \right)^{\frac{s_i^0 + s_i^1}{2}} \prod_{i \in S_D^0} \left(\frac{\tilde{p}_i^1}{p_i^0} \right)^{\frac{s_i^0}{2}} \prod_{i \in S_N^1} \left(\frac{p_i^1}{\tilde{p}_i^0} \right)^{\frac{s_i^1}{2}}. \qquad (5)$$

224 Triplett (2004) refers to this approach as the characteristics price index method since it uses the implicit characteristics prices (the regression coefficients from the hedonic function) in a conventional index number formula. Silver and Heravi (2004) call it the hedonic imputation method. This name may be confusing, so I termed it *full* hedonic imputation.

225 This idea is not new, of course. Triplett (2004) notes the following on the (partial) hedonic imputation method: “Where matched model comparisons are possible, they are used. Where they are not possible, a hedonic imputation is made for the item replacement. Hedonic imputation methods make maximum use of observed data, and minimum use of imputation, thereby minimizing estimation variance. The hedonic imputation method was employed in the hedonic computer indexes introduced into the U.S. national accounts in 1985”. Diewert (2003) also suggests matching items where possible and using hedonic regressions to impute the ‘missing prices’.

I prefer P_{PHIT} as the (geometric) target index because it is a generalized superlative index that restricts hedonic imputation to the 'missing prices'.[226]

Diewert (2003) argues that the residuals from a logarithmic hedonic model are less likely to be heteroskedastic than those from a linear model. Most empirical studies seem to prefer the logarithmic specification as well. For period t (t= 0,1) the semi-log (log-linear) model regression can be expressed as

$$\ln p_i^t = \alpha^t + \sum_{k=1}^{K} \beta_k^t z_{ik} + \varepsilon_i^t , \qquad (6)$$

where z_{ik} is the k-th characteristic of item i and β_k^t the corresponding parameter; the errors ε_i^t are assumed to be independently distributed with expected values of zero and constant variances. Least squares estimation of (6) on the data from each period separately produces regression coefficients $\tilde{\alpha}^t$ and β_k^t and predicted prices $\tilde{p}_i^t = \exp(\tilde{\alpha}^t + \sum_{k=1}^{K} \beta_k^t z_{ik})$.[227]

The time dummy index

The second main approach to estimating hedonic price indexes is the time dummy variable method. Here the cross section data from period 0 and period 1 are *pooled*. The semi-log hedonic model reads

$$\ln p_i^t = \alpha + \delta D_i^t + \sum_{k=1}^{K} \beta_k z_{ik} + \varepsilon_i^t , \qquad (7)$$

where D_i^t is a dummy variable that takes on the value of 1 if the observation comes from period 1 and 0 otherwise. The errors ε_i^t are now assumed to be similarly and independently distributed in both periods, which may be rather restrictive. Model (7) assumes that each of the characteristics parameters is the same across the two time periods compared, i.e. $\beta_k^0 = \beta_k^1 = \beta_k$ for k=1,….,K by assumption. As (7) controls for changes in the quality characteristics, the exponent of the estimated time dummy coefficient $\hat{\delta}$ directly produces a quality-adjusted measure of price change. The predicted prices in period 0 and period 1 are denoted $\hat{p}_i^0 = \exp(\hat{\alpha} + \sum_{k=1}^{K} \beta_k z_{ik})$ and $\hat{p}_i^1 = \exp(\hat{\alpha} + \delta + \sum_{k=1}^{K} \beta_k z_{ik})$. Note that $\hat{p}_i^1 / \hat{p}_i^0 = \exp \hat{\delta}$ for all i.

We may ask the question under what circumstances $\exp \hat{\delta}$ can be interpreted as an aggregate price index, based on a conventional index number formula. Silver (2003) criticises the use of Ordinary Least Squares (OLS) to estimate (7). The observations should be weighted according to their economic importance. Van der Grient and De Haan (2003) present a general decomposition of the Weighted Least Squares (WLS) time dummy index which I will repeat here and extend for further analysis. Let w_i^0 and w_i^1 denote the regression weights for $i \in S^0$ and $i \in S^1$, respectively. That is, each observation in period t counts w_i^t times in the regression. Because a constant term is included in (7), the (weighted) residuals $u_i^t = \ln p_i^t - \ln \hat{p}_i^t = \ln(p_i^t / \hat{p}_i^t)$ sum to zero in each period: $\sum_{i \in S^0} w_i^0 \ln(p_i^0 / \hat{p}_i^0) = \sum_{i \in S^1} w_i^1 \ln(p_i^1 / \hat{p}_i^1) = 0$. By exponentiating both sums we obtain the following relation:

[226] See De Haan (2004a) for a *double* imputation hedonic geometric index in which period 0 and period 1 prices are imputed for both new and disappearing items (but not for matched items). De Haan (2002) calls the partial imputation Fisher index a generalized Fisher index.

[227] The predictors are not unbiased estimators of the actual prices. Van Dalen and Bode (2004) evaluate the biases in various hedonic indexes due to the use of a logarithmic hedonic model.

$$\prod_{i\in S_M}\left(\frac{p_i^0}{\hat{p}_i^0}\right)^{w_i^0}\prod_{i\in S_D^0}\left(\frac{p_i^0}{\hat{p}_i^0}\right)^{w_i^0}=\prod_{i\in S_M}\left(\frac{p_i^1}{\hat{p}_i^1}\right)^{w_i^1}\prod_{i\in S_N^1}\left(\frac{p_i^1}{\hat{p}_i^1}\right)^{w_i^1}(=1). \qquad (8)$$

Substituting $\hat{p}_i^1/\hat{p}_i^0=\exp\hat{\delta}$ for $i\in S_M$ and defining $w_M^t=\sum_{i\in S_M}w_i^t$ leads to a decomposition of the WLS time dummy index:

$$P_{TD}=\exp\hat{\delta}=\prod_{i\in S_M}\left(\frac{p_i^1}{p_i^0}\right)^{\frac{w_i^1}{w_M^1}}\prod_{i\in S_M}\left(\frac{p_i^0}{\hat{p}_i^0}\right)^{\frac{w_i^1-w_i^0}{w_M^1}}\prod_{i\in S_D^0}\left(\frac{p_i^0}{\hat{p}_i^0}\right)^{\frac{-w_i^0}{w_M^1}}\prod_{i\in S_N^1}\left(\frac{p_i^1}{\hat{p}_i^1}\right)^{\frac{w_i^1}{w_M^1}}. \qquad (9)$$

Van der Grient and De Haan (2003) formulate two requirements equation (9) should satisfy. The *first requirement* is that the resulting index should be based on observed prices when there are no new or disappearing items. Quality changes do not occur in that case, although the quality mix may change because of changes in the quantities sold, and we want the outcome to be independent of the set of characteristics. The second factor of (9) contains the period 0 residuals of the matched items and usually differs from unity. Hence, the first requirement will not be met. We therefore impose the restriction $w_i^0=w_i^1=w_i$ (independent of time) for $i\in S_M$, which assures that the time dummy index equals the matched-item index $P_M=\prod_{i\in S_M}(p_i^1/p_i^0)^{w_i/w_M}$, where $w_M=\sum_{i\in S_M}w_i$, when there are only matched items.[228] This yields

$$P_{TD(R)}=\prod_{i\in S_M}\left(\frac{p_i^1}{p_i^0}\right)^{\frac{w_i}{w_M}}\prod_{i\in S_D^0}\left(\frac{p_i^0}{\hat{p}_i^0}\right)^{\frac{-w_i^0}{w_M}}\prod_{i\in S_N^1}\left(\frac{p_i^1}{\hat{p}_i^1}\right)^{\frac{w_i^1}{w_M}}. \qquad (10)$$

Equation (10) shows an important property of the time dummy index satisfying the above restriction: it can be viewed as an index where all observed prices, including those of the unmatched items, are kept and not replaced by their predicted values. This is in line with the partial hedonic imputation indexes defined in section 2. What I did not realize before was that, by using $\hat{p}_i^1/\hat{p}_i^0=\exp\hat{\delta}$ for $i\in S_D^0$ and $i\in S_N^1$, equation (10) can be rewritten as

$$P_{TD(R)}=\left[\prod_{i\in S_M}\left(\frac{p_i^1}{p_i^0}\right)^{w_i}\prod_{i\in S_D^0}\left(\frac{\hat{p}_i^1}{p_i^0}\right)^{w_i^0}\prod_{i\in S_N^1}\left(\frac{p_i^1}{\hat{p}_i^0}\right)^{w_i^1}\right]^{\frac{1}{w_M+w_D^0+w_N^1}}, \qquad (11)$$

where $w_D^0=\sum_{i\in S_D^0}w_i^0$ and $w_N^1=\sum_{i\in S_N^1}w_i^1$. Notice that the term between square brackets has the same structure as P_{PHIT} given by (5). The time dummy index (11) can be rewritten as a weighted geometric average of the matched-model index P_M defined above and hedonic imputation indexes for disappearing and new items:

$$P_{TD(R)}=\left[P_M\right]^{\frac{w_M}{w_M+w_D^0+w_N^1}}\left[\hat{P}_D\right]^{\frac{w_D^0}{w_M+w_D^0+w_N^1}}\left[\hat{P}_N\right]^{\frac{w_N^1}{w_M+w_D^0+w_N^1}}, \qquad (11')$$

where $\hat{P}_D=\prod_{i\in S_D^0}(\hat{p}_i^1/p_i^0)^{w_i^0/w_D^0}$ and $\hat{P}_N=\prod_{i\in S_N^1}(p_i^1/\hat{p}_i^0)^{w_i^1/w_N^1}$.

[228] This restriction does not hold if there is no weighting of the observations involved (OLS), unless the number of observations is constant over time, e.g. under the fixed-size sampling schemes applied by most statistical agencies. Yet the use of OLS always satisfies the first requirement: it produces the unweighted geometric index when all items are matched.

A *second requirement* is that the resulting index number formula can be defended on theoretical grounds. This implies that the price relatives must somehow be weighted by expenditure shares.[229] The use of OLS is no longer an option unless we are dealing with fixed-size sets of items sampled proportional to expenditure (De Haan, 2003). The choice of regression weights is explored in section 4.

Choice of weights

For the matched items we have expenditure shares of both period 0 and period 1. For reasons of symmetry their unweighted arithmetic average is a natural choice, which also meets the first requirement. This choice has been suggested by Diewert (2003); the time dummy index would then equal the superlative Törnqvist index when there are only matched items. Substituting $w_i = (s_i^0 + s_i^1)/2$ for $i \in S_M$ into (11) yields the (pseudo) generalized Törnqvist index

$$\widetilde{P}_{TD(GT)} = \left[\prod_{i \in S_M} \left(\frac{p_i^1}{p_i^0} \right)^{\frac{s_i^0 + s_i^1}{2}} \prod_{i \in S_D^0} \left(\frac{\hat{p}_i^1}{p_i^0} \right)^{w_i^0} \prod_{i \in S_N^1} \left(\frac{p_i^1}{\hat{p}_i^0} \right)^{w_i^1} \right]^{\frac{1}{\frac{s_M^0 + s_M^1}{2} + w_D^0 + w_N^1}}, \qquad (12)$$

using $w_M = (s_M^0 + s_M^1)/2$, where $s_M^0 = \sum_{i \in S_M} s_i^0$ and $s_M^1 = \sum_{i \in S_M} s_i^1$ denote the matched items expenditure shares in period 0 and period 1, respectively.

For the unmatched items the choice of weights is less obvious. I propose to take half the expenditure shares in the periods they are available, i.e. $w_i^0 = s_i^0 / 2$ for $i \in S_D^0$ and $w_i^1 = s_i^1 / 2$ for $i \in S_N^1$. Substituting those weights into (12) gives

$$P_{TD(H)} = \prod_{i \in S_M} \left(\frac{p_i^1}{p_i^0} \right)^{\frac{s_i^0 + s_i^1}{2}} \prod_{i \in S_D^0} \left(\frac{\hat{p}_i^1}{p_i^0} \right)^{\frac{s_i^0}{2}} \prod_{i \in S_N^1} \left(\frac{p_i^1}{\hat{p}_i^0} \right)^{\frac{s_i^1}{2}}, \qquad (13)$$

since now $w_D^0 = (1 - s_M^0)/2$ and $w_N^1 = (1 - s_M^1)/2$, so that $w_M + w_D^0 + w_N^1 = 1$. Expression (13) is similar to (5). Thus, using the proposed weights, the time dummy index is a *special case* of the partial hedonic imputation Törnqvist price index (5) in which the 'missing prices' are automatically imputed according to the time dummy variable model (7). Notice that the regression weights are identical to the weights used to aggregate the price relatives. De Haan (2004a) also applies (7) in a hedonic imputation approach and refers to this as an *indirect* time dummy approach. In the special case discussed here the direct and indirect approaches coincide. This seems a desirable property. Diewert (2003), on the other hand, suggests taking the full expenditure shares of the unmatched items as regression weights. His suggestion has been followed by Silver and Heravi (2004) and also by Van der Grient and De Haan (2003), Van der Grient (2004), and De Haan (2004b). It is mentioned in the new CPI manual (ILO, 2004) as well.[230] Substituting $w_i^0 = s_i^0$ for $i \in S_D^0$ and $w_i^1 = s_i^1$ for $i \in S_N^1$ into (12) yields

[229] Various authors, for example Silver and Heravi (2002) and Van Mulligen (2003), have used the quantities sold in the respective periods as weights in the WLS estimation of time dummy indexes. This procedure violates both requirements and should be adviced against.

[230] See Silver (2003) for a general discussion on the use of weights.

$$P_{TD(D)} = \left[\prod_{i \in S_M} \left(\frac{p_i^1}{p_i^0} \right)^{\frac{s_i^0 + s_i^1}{2}} \right]^{\frac{2}{4 - s_M^0 - s_M^1}} \left[\prod_{i \in S_D^0} \left(\frac{\hat{p}_i^1}{p_i^0} \right)^{\frac{s_i^0}{2}} \right]^{\frac{4}{4 - s_M^0 - s_M^1}} \left[\prod_{i \in S_N^1} \left(\frac{p_i^1}{\hat{p}_i^0} \right)^{\frac{s_i^1}{2}} \right]^{\frac{4}{4 - s_M^0 - s_M^1}} . \quad (14)$$

I find it difficult to justify this choice of weights since expression (14) has no clear interpretation as a generalized Törnqvist index. The weight of the matched items is understated (because $s_M^0 + s_M^1 < 2$) and the weights of new and disappearing items accordingly overstated. The overall impact also depends on the imputed prices of the unmatched items which will naturally differ using different weights. However, if the model specification is satisfactory, we may expect this effect to be relatively small compared to the effect of the 'wrong' weights with which the price relatives are implicitly aggregated.

Now that we have determined the 'optimal' regression weights, various expressions or decompositions of the time dummy index can be derived to provide some further insight. For example, equation (13) can be rewritten in the form of (11'), yielding

$$P_{TD(H)} = \left[\prod_{i \in S_M} \left(\frac{p_i^1}{p_i^0} \right)^{\frac{s_i^0 + s_i^1}{s_M^0 + s_M^1}} \right]^{\frac{s_M^0 + s_M^1}{2}} \left[P_{DGL} \right]^{\frac{1 - s_M^0}{2}} \left[P_{NGP} \right]^{\frac{1 - s_M^1}{2}} , \quad (15)$$

where $P_{DGL} = \prod_{i \in S_D^0} (\hat{p}_i^1 / p_i^0)^{s_i^0/(1 - s_M^0)}$ and $P_{NGP} = \prod_{i \in S_N^1} (p_i^1 / \hat{p}_i^0)^{s_i^1/(1 - s_M^1)}$ are the hedonic imputation geometric Laspeyres index and geometric Paasche index for disappearing items and new items, where the imputations are based on model (7).

A second and perhaps more convenient decomposition results from substituting the proposed weights into equation (10). We then obtain

$$P_{TD(H)} = \prod_{i \in S_M} \left(\frac{p_i^1}{p_i^0} \right)^{\frac{s_i^0 + s_i^1}{s_M^0 + s_M^1}} \frac{\left[\exp \bar{u}_N^1 \right]^{\frac{1 - s_M^1}{s_M^0 + s_M^1}}}{\left[\exp \bar{u}_D^0 \right]^{\frac{1 - s_M^0}{s_M^0 + s_M^1}}} , \quad (16)$$

in which $\bar{u}_N^1 = \sum_{i \in S_N^1} s_i^1 u_i^1 / \sum_{i \in S_N^1} s_i^1$ and $\bar{u}_D^0 = \sum_{i \in S_D^0} s_i^0 u_i^0 / \sum_{i \in S_D^0} s_i^0$ denote the expenditure share weighted average residuals of new and disappearing items. Under perfect competition the so-called law of one quality-adjusted price predicts that there are no items whose prices are relatively high or low given their characteristics: there should be no 'unusual' prices. In this case $\bar{u}_D^0$ and $\bar{u}_N^1$ in (16) approximate zero so that the time dummy index will be approximately equal to the matched-item index $\prod_{i \in S_M} (p_i^1 / p_i^0)^{(s_i^0 + s_i^1)/(s_M^0 + s_M^1)}$. But under imperfect competition there are economic reasons to expect that such unusual prices exist (Silver and Heravi, 2002; Triplett, 2004).

Econometric considerations

Because the time dummy index (13) is a special case of the broader class of (partial) imputation Törnqvist indexes, one could argue that it would be better to estimate the 'missing prices' *explicitly*, based on econometric arguments, rather than implicitly as the time dummy index does. From an econometric point of view there are again the issues of pooling data and weighting. Pooling data is of course not necessary for an imputation approach. The assumption of constant parameters is rather stringent, and estimating (log-linear) hedonic models from data of each period separately would in principle be best. Indeed, this has been suggested in section 2 and it is also what Silver and Heravi

(2004) do in the estimation of the full hedonic imputation indexes. But pooling data has two advantages. First, there is an expected gain in efficiency by saving degrees of freedom, which can be particularly helpful when the sample size is relatively small. Second, it makes the estimation of period 0 prices possible for new items having characteristics that did not exist at that time. Moreover, although the underlying assumption is restrictive in theory, many if not most empirical studies find the parameters to be fairly stable in the short run.

Let us compare the proposed time dummy index with the partial hedonic imputation Törnqvist index estimated on data from each period separately. Dividing (13) by (5) and substituting the expressions for the unmatched items' predicted prices gives

$$\frac{P_{TD(H)}}{P_{PHIT}} = \frac{\left[\exp\left\{\hat{\alpha} + \hat{\delta} - \tilde{\alpha}^1 + \sum_{k=1}^{K}\left(\hat{\beta}_k - \tilde{\beta}_k^1\right)\bar{z}_{Dk}^0\right\}\right]^{\frac{1-s_M^0}{2}}}{\left[\exp\left\{\hat{\alpha} - \tilde{\alpha}^0 + \sum_{k=1}^{K}\left(\tilde{\beta}_k^0 - \hat{\beta}_k\right)\bar{z}_{Nk}^1\right\}\right]^{\frac{1-s_M^1}{2}}}, \tag{17}$$

where $\bar{z}_{Dk}^0 = \sum_{i \in S_D^0} s_i^0 z_{ik} / \sum_{i \in S_D^0} s_i^0$ and $\bar{z}_{Nk}^1 = \sum_{i \in S_N^1} s_i^1 z_{ik} / \sum_{i \in S_N^1} s_i^1$ denote the expenditure-weighted average characteristics of disappearing and new items. If the assumption of constant characteristics parameters (approximately) holds, we expect to find $\hat{\beta}_k \approx \tilde{\beta}_k^0 \approx \tilde{\beta}_k^1$. For simplicity I assume that $s_M^1 \approx s_M^0$. Now (17) reduces to

$$\frac{P_{TD(H)}}{P_{PHIT}} \approx \left[\exp\left\{\hat{\delta} - \left(\tilde{\alpha}^1 - \tilde{\alpha}^0\right)\right\}\right]^{\frac{1-s_M^0}{2}}. \tag{18}$$

Expression (18) does not seem helpful at first sight because the time dummy index $P_{TD(H)} = \exp\hat{\delta}$ appears on the right-hand side. However, it underlines the fact that, from an imputation point of view, δ in (7) is viewed as a *common shift parameter* needed to justify the pooling of data from two data sets, pertaining to two periods, under the assumptions of constant characteristics parameters and independently and identically distributed errors. Stated otherwise, under these assumptions $\hat{\delta}$ is merely an estimator of $\alpha^1 - \alpha^0$ and $P_{TD(H)}$ is unlikely to differ much from P_{PHIT}, provided that $s_M^1 \approx s_M^0$. $P_{TD(H)}$ is the more efficient estimator due to the pooling of data but it will be somewhat biased when the parameters are not constant over time.[231]

The choice of weights has to do with *representativity*. The chosen weights – or any other set of weights for that matter – may not be optimal from an econometric point of view. Econometrics textbooks recommend the use of WLS to reduce the variance of the coefficients if there is evidence of heteroskedasticity. On the other hand, if the errors are homoskedastic (as assumed), OLS is preferred in general and WLS could introduce some loss in efficiency.[232] That would partly counteract the efficiency gain resulting from the pooling of data. The problem is of course that by using OLS to estimate model (7) on the pooled data we are 'running around in circles' because the similarity between the (direct) time dummy approach and the imputation Törnqvist index (the 'indirect' time dummy approach) no longer holds. To avoid interpretation problems, the proposed

[231] From a statistical point of view the mean square error (i.e. the sum of the variance and the squared bias) would be the natural criterion to choose between alternative estimators. Hence even if the parameters were slightly changing over time, the time dummy index could be the 'best' index.

[232] If the weights are exogenous, least squares estimators remain unbiased. The weights are not truly exogenous since they contain the dependent variable (price) of the model, which makes them stochastic variables. This could introduce some bias in the WLS estimators.

time dummy index might well be the only serious option if one wishes to pool cross-section data to estimate a quality-adjusted price index using a logarithmic hedonic model.
The time dummy index essentially imputes 'missing prices'. As with any imputation method, we have to act as if the data had been generated by the underlying model. In this particular case it is assumed that the (log of) prices are generated according to model (7). One issue I mentioned earlier (De Haan, 2003, 2004a; Van der Grient and De Haan, 2003) keeps bothering me. Suppose we find empirically that the weighted average residuals of the unmatched items in expression (16) differ substantially from zero. If the number of unmatched items is sufficiently large, this could be taken as evidence of systematic patterns in the unmatched items' residuals. This suggests that the assumption of a zero expected value of the errors is violated, so that estimating (7) by least squares regression produces biased parameter estimators. Isn't there an inconsistency in assuming that the (log of) prices are generated according to model (7) and at the same time finding that the observed (log of) prices of the unmatched items differ systematically from their model-based predictions? De Haan (2003) has argued that the direct time dummy approach cannot cope with such systematic price effects and suggested to incorporate dummy variables for the unmatched items. But that is beyond the scope of this paper.

Empirical illustration: PCs

The effect of the use of the 'wrong' regression weights on the time dummy index is an empirical matter. This section presents empirical evidence for PCs using scanner data from market research company GfK. The monthly data on prices (unit values), quantities sold and product characteristics cover the entire Dutch consumer market during 1999-2001. A comprehensive data description can be found in Van der Grient (2004).
Individual PC models are usually identified by model numbers or, as in our case, by bar codes assigned by the manufacturers. A specific model sold in different outlets does not necessarily yield the same utility to the consumer. PCs having identical bar codes but sold in different types of outlets are therefore treated as different items. Six outlet types are distinguished in the data set, e.g. chains, buying combinations, independents, and specialized stores. The latter are responsible for the greater part of PC sales. The market for PCs is an extremely dynamic one. On average only two thirds of the items are matched between adjacent months. Because the best selling items have a relatively long 'lifetime', the average monthly expenditure share of the matched items is somewhat higher and amounts to 0.8. To give an example: only 5% of the items stay on the market for more than a year, but they account for 30% of total turnover.
New PC models often have new features, e.g. a 'faster' processor. The time dummy index deals with this by estimating and imputing base period prices for such models. This poses theoretical problems, however, if the new technology was not available in the base period. This will be particularly problematic when the base period and the current period are far apart. Apart from that, it seems undesirable to maintain the assumption of constant parameters for a long time span in highly dynamic markets. High-frequency chaining is inevitable here. So unlike Silver and Heravi (2004), who compute direct indexes, I will estimate monthly chained indexes. This may not be without problems either because PC sales clearly exhibit a seasonal pattern.[233] As is well known, seasonality might create drift in monthly-chained indexes.

[233] This is true at the aggregate level. However, it is a bit difficult to speak of seasonality at the item level since 95% of the items are available on the market for less than a year.

The semi-log time dummy model includes dummy variables for both brand and type of outlet. See De Haan (2004b) for some (OLS) regression results. Most coefficients differ significantly from zero at the 5%-level and their signs are in accordance with a priori expectations. No evidence of heteroskedasticity was found. The adjusted R^2 does not exceed 0.71, which is a bit disappointing. This may be due to the fact that only a limited number of technical or performance characteristics are included in the data set. Unobserved but relevant characteristics could give rise to omitted variables bias. The model includes three continuous performance characteristics: processor speed in MHz, storage capacity of the hard disk in MB and memory capacity in MB. In addition there are three qualitative characteristics (apart from brand and type of outlet): type of processor, the presence of a monitor and the availability of a USB port.

Figure 1 depicts four series of monthly-chained price indexes. Two series pertain to WLS time dummy hedonic indexes, one estimated with the proposed weights and another with the weights suggested by Diewert (2003). The remaining two series are added for comparison and pertain to the OLS time dummy index and the matched-item Törnqvist index $P_{MT} = \prod_{i \in S_M} (p_i^1 / p_i^0)^{(s_{iM}^0 + s_{iM}^1)/2}$, where s_{iM}^t is the share of *i* in the period *t* expenditures of the matched-items. All indexes have similar trends and point towards large declines in quality-adjusted PC prices, as expected. The time dummy price index estimated with the 'wrong' weights falls at a faster rate than the proposed time dummy index, although the difference is very small compared to the huge price decrease itself. In January 2002 the index numbers are 31.84 and 34.57. This result indicates that new and disappearing items had a combined downward effect on the quality-adjusted index.

The use of the matched-item Törnqvist index would have led to an annual average upward bias of 1.5%-points during the three-year period. Surprisingly, the OLS time dummy index, despite being an unweighted index, performs quite well.

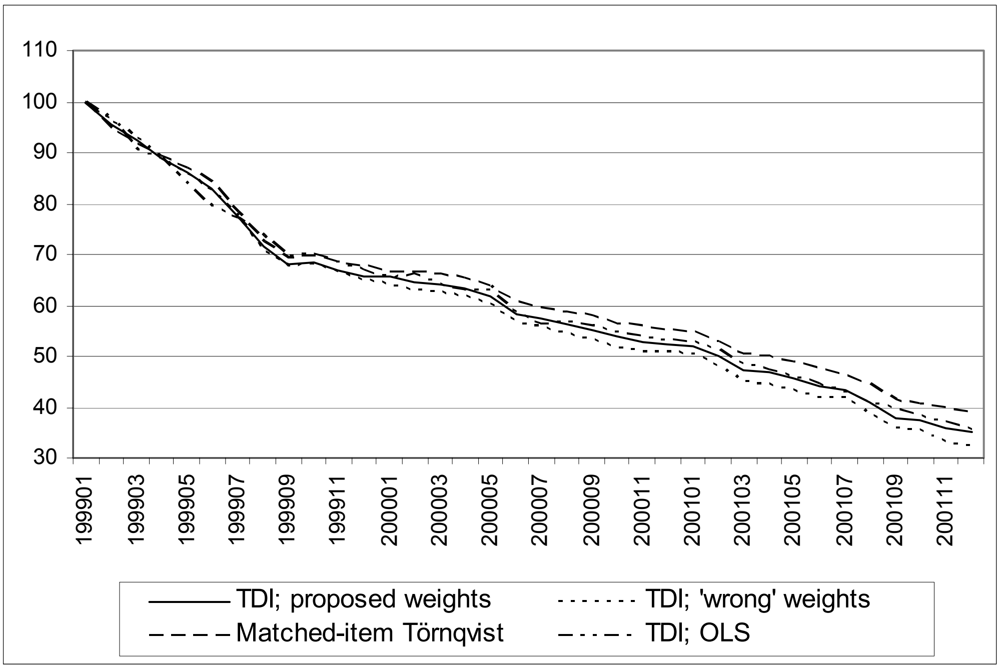


Figure 1. Monthly chained price indexes for PCs (January 1999= 100)

Conclusion

Using a simple framework this paper argues that the time dummy hedonic index can be interpreted as a special case of the hedonic imputation Törnqvist price index if the regression weights are properly chosen. The weights for the unmatched (new and disappearing) items suggested in the recent literature on hedonics are twice as large as they should be if the aim is to estimate a generalized Törnqvist index. The choice of weights most likely affects the index for high-turnover goods like computers. An empirical illustration on scanner data for the Netherlands indicated that the use of the 'wrong' weights would have underestimated the monthly chained time dummy index for PCs during 1999 – 2001 by some 2.7 index points.

References

Dalen, J. van, and B. Bode, 2004, Estimation Biases in Quality-Adjusted Hedonic Price Indices, *Paper presented at the SSHRC International Conference on Index Number Theory and the Measurement of Prices and Productivity, Vancouver*, June 30 – July 3, 2004.

Diewert, W.E., 2003, Hedonic Regressions: A Review of Some Unresolved Issues, *Proceedings of the Seventh Meeting of the International Working Group on Price Indices*, (ed. Th. Lacroix), pp. 71-109. Paris: INSEE.

Grient, H.A. van der, 2004, Scanner Data on Durable Goods: Market Dynamics and Hedonic Time Dummy Indexes, *Discussion paper no. 04011*, Statistics Netherlands, Voorburg.

Grient, H.A. van der, and J. de Haan, 2003, An Almost Ideal Hedonic Price Index for Televisions, *Proceedings of the Seventh Meeting of the International Working Group on Price Indices*, (ed. Th. Lacroix), pp. 111-121. Paris: INSEE.

Haan, J. de, 2002, Generalised Fisher Price Indexes and the Use of Scanner Data in the CPI, *Journal of Official Statistics*, 18, vol. 1, 61-85.

Haan, J. de, 2004a, Direct and Indirect Time Dummy Approaches to Hedonic Price Measurement, *Journal of Economic and Social Measurement* (forthcoming).

Haan, J. de, 2004b, Estimating Quality-Adjusted Unit Value Indexes: Evidence from Scanner Data, *Paper presented at the SSHRC International Conference on Index Number Theory and the Measurement of Prices and Productivity*, Vancouver, June 30 – July 3, 2004.

International Labour Office (ILO), 2004, *Consumer Price Index Manual; Theory and Practice*. Geneva: ILO Publications.

Mulligen, P.H. van, 2003, *Quality Aspects in Price Indices and International Comparisons: Applications of the Hedonic Method*, Ph.D. Dissertation, University of Groningen. Voorburg: Statistics Netherlands.

Silver, M., 2003, The Use of Weights in Hedonic Regressions: The Measurement of Quality-Adjusted Price Changes, *Proceedings of the Seventh Meeting of the International Working Group on Price Indices*, (ed. Th. Lacroix), pp. 135-147. Paris: INSEE.

Silver, M. and S. Heravi, 2002, Why the CPI Matched Models Method May Fail Us: Results from an Hedonic and Matched Experiment Using Scanner Data, *European Central Bank (ECB) Working Paper No. 144*, ECB, Frankfurt.

Silver, M. and S. Heravi, 2004, The Difference between Hedonic Imputation Indexes and Time Dummy Hedonic Indexes for Desktop PCs, *Paper presented at the CRIW Conference on Price Index Concepts and Measurement*, Vancouver, June 28-29, 2004.

Triplett, J.E., 2004, *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes: Special Application to Information Technology Products*. Paris: OECD. Available at www.brook.edu/views/papers/triplett/200407_book.htm.

Table 1, Number of matched and unmatched models and matched models and hedonic indexes

	Number of matched models	Number of unmatched old models in January of the comparison	Number of unmatched new models in the current month of the comparison	Matched models Törnqvist index	**HI** Törnqvist index	**DTH** Törnqvist index
	Figures are for comparisons between January and each current month			Fixed base, January 1998=1.000		
February	584	86	104	0.945	0.939	0.939
March	577	93	181	0.887	0.877	0.877
April	346	324	191	0.825	0.802	0.802
May	315	355	227	0.767	0.738	0.737
June	297	373	265	0.736	0.656	0.658
July	282	388	301	0.765	0.618	0.623
August	276	394	351	0.769	0.616	0.626
September	247	423	382	0.777	0.599	0.629
October	193	477	402	0.772	0.518	0.551
November	164	506	435	0.765	0.458	0.495
December	161	509	436	0.764	0.452	0.496

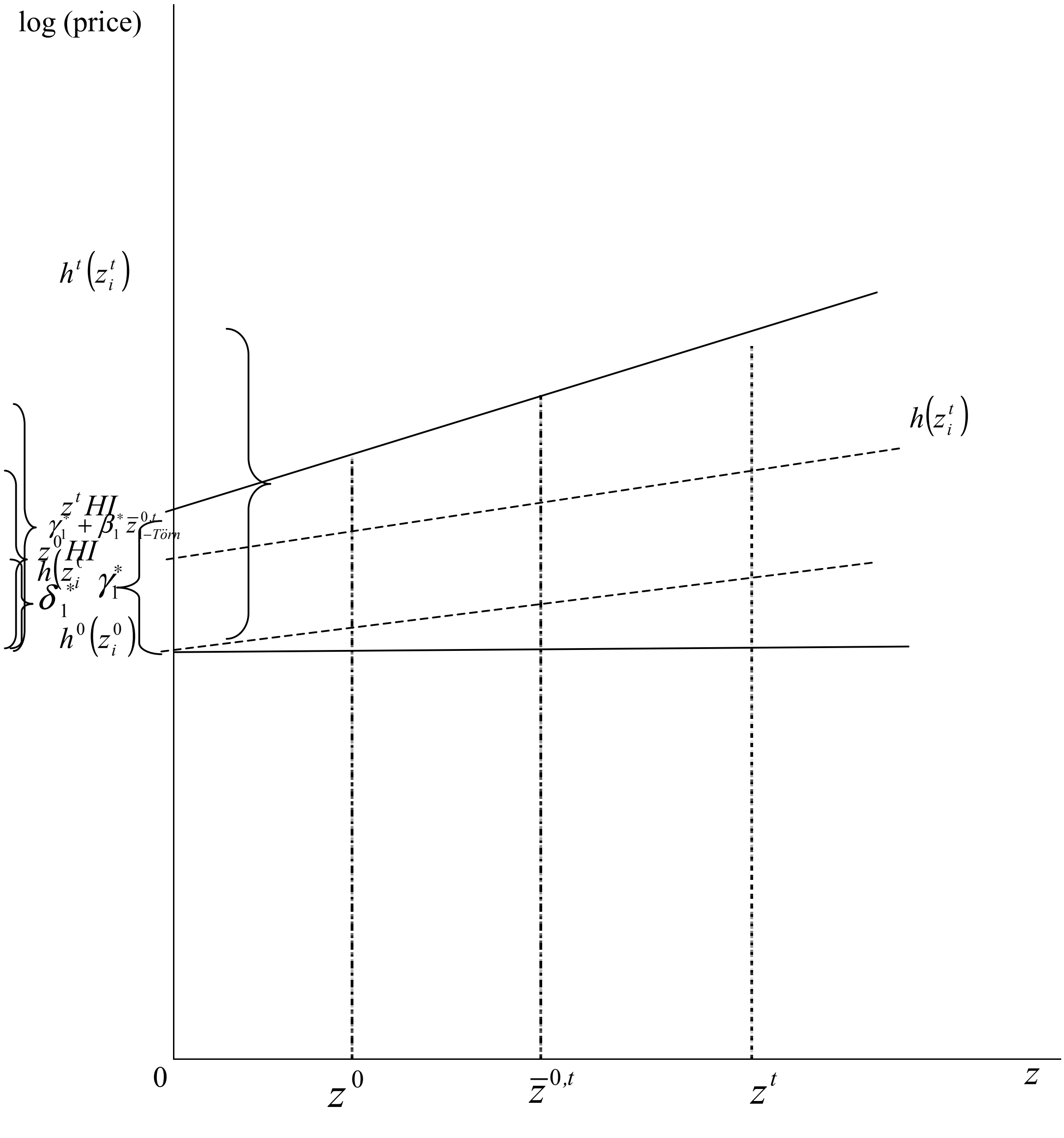
Table 2, Decomposition of differences between HI and Törnqvist indexes for desktop PCs, January 1998=1.000*

	HI ÷ DTH Törnqvist indexes (1)	$(\beta_1^{*t}-\beta_1^{*0})$ (2)	$(\lambda_1^{*}-\bar{z}_{1-\text{Törn}}^{0,t})$ (3)	$(\lambda_1^{*}-\bar{z}_{1-\text{Törn}}^{0,t})\times(\beta_1^{*t}-\beta_1^{*0})$ (4)	$\bar{z}_{1-\text{Törn}}^{0,t}$ (5)	λ_1^{*} (6)	$(\bar{z}_1^{t}-\bar{z}_1^{0,t})$ (7)	$\text{cov}(D_{1i}^{0,t},z_{1i}^{0,t})$ (8)	$(2-\sum_{i\in S(t}s_i^{0,}$ (9)†
Fixed base, January 1998=1.000									
February	0.999999	-0.00042	0.00292	-0.0000012	208.0586	208.0616	-0.15937	1193.948	0.997
March	1.000001	-0.0000021	-0.05901	0.0000013	211.0463	210.9873	0.961408	1323.364	0.9828
April	1.000121	0.000588	0.20158	0.0001186	219.3599	219.5615	7.047697	1952.174	0.99332
May	0.998721	0.00069	-1.86554	-0.0012871	225.0111	223.1456	11.53865	2911.7	0.95714
June	1.002505	-0.00033	-7.59551	0.0025079	231.0368	223.4413	16.03635	4207.419	0.87911
July	1.008229	-0.00078	-10.56785	0.0082167	236.5226	225.9548	18.31796	4804.071	0.83007
August	1.015499	-0.00081	-19.01683	0.0154110	244.4779	225.4611	23.22285	6481.596	0.79600
September	1.050295	-0.00172	-28.49892	0.0491427	250.8653	222.3664	26.35361	8272.112	0.75408
October	1.063765	-0.00154	-40.14595	0.0619327	262.9021	222.7562	34.28005	10189.09	0.73220
November	1.079882	-0.00166	-46.48257	0.0770125	269.5905	223.1079	37.8419	11185.42	0.71721
December	1.097709	-0.00177	-52.66507	0.0933989	273.4374	220.7723	39.68194	12243.07	0.7073

- All statistics are weighted using the weights commensurate with a Törnqvist index as outlined in the text following equation (7).

† The expression $(2-\sum_{i\in S(t}s_i^{0,t})$ replaces $(N-N')$ in (17) since the statistics are weighted. The difference is that between the sum of the shares over the two periods, i.e. 2, and the sum of that share arising from observations in period t. With new models taking on an increased share of sales the statistic falls.

Figure 1


log (price)
$h^t(z_i^t)$
$h(z_i^t)$
$h^0(z_i^0)$
γ_1^*
δ_1^*
0
z^0
$\bar{z}^{0,t}$
z^t
z

Kari Manninen: The Effects of Quality Adjustment Methods on Price Indices for Digital Cameras

Abstract

In this paper we test different hedonic and conventional quality adjustment methods in an uniform, but somewhat unconventional, descriptive framework. The main aim is to address questions on hedonic quality adjustment methods and their robustness in index compilation and to give an empirical example with digital camera prices. We will show how conventional quality adjusting methods may be treated parallel with hedonic ones and how these methods may be evaluated similarly with regression based methods. Contrary to structural models that many hedonic quality adjusted price indices are based on, the hedonic models in this paper are all used as forecast models which, we believe, add to the robustness and practical utility of hedonics as a tool for statistical agencies using quality adjustment.

The empirical part of the paper is based on findings from a quarterly digital camera database including some 1,200 prices from over 250 different digital camera models over the years 1998 to 2002. The main findings indicate that, in an aggregate context, such as price index, relatively simple hedonic models may be sufficient for accurate quality controlling even in high technology products. Further, if compared with a matched model framework, the collection of characteristics data for hedonics may not need to exceed the precision already needed in the matched model. This suggests that it may be feasible to use hedonic indices even in high frequency index compilation. To validate this, we claim that additional cross sectional explanatory power from a set of added quality characteristics in hedonic models have only marginal longitudinal effects in the index series.

Contents:

I Introduction

Constructing quality adjusted price indices for high technology goods have been discussed for years. In this article we take part in the discussion and use digital cameras as an example.

There is a sound mutual understanding among the academia and statistical agencies that, in volume and piece measures, changes in quality of goods and services have to be accounted for. Various regression based methods, or hedonic methods, have been developed and used in a variety of ways to control for quality in index number calculations. Even though theoretical foundations for use of different methods have been made, most notably Fixler and Zieschang (1992), Feenstra (1995), and Diewert (2001), the spectrum of methods still blurs the answer whether the results are a figment of the technique that is used. Especially with high frequency indices such as the CPIs and PPIs, the theoretical justifications are hard to make use operationally. This paper tests the sensitivity of indices for digital cameras to the choice of these techniques. The main finding is that all reasonable methods give approximately the same answer.

In the second section we introduce, within a unified estimation approach, six seemingly different methods; grand unit-value, class unit-value, matched model, time dummy pooled regression, time dummy 2-period regression, and full hedonic regression. We compare the results of applying these methods to the digital camera data. Somewhat similar comparative studies for different data sets have been made at least by Silver (1999), Moulton et al. (1999), Kokoski et al. (1999) and, Silver et. Heravi (2000). These studies mainly focus on differences between matched model and hedonic methods. Few studies are made on differences that result from applying different hedonic methods.

Next, we discuss the advantages and shortcomings of each method. We discuss the implicit assumptions each method imposes and weight how justified these assumptions may be.

In section four we lay out the data for digital cameras and discuss the sample frame and some reservations. In section five present the results for the six different methods introduced earlier. The presentation is the resulting quality adjusted price indices.

Finally, we give conclusions and propose some additional topics to be investigated.

The aim of this exercise is twofold. First, we want to assess the effect of quality adjustment on the price index and to show how the model specification affects the quality adjustment and the index. Second, we want to study the implied accuracy and robustness of the hedonic index and compare the indices based on different types of models.

All indices are treated as changes between two consecutive periods. The index strategy is to chain the index over the periods. In some cases, an equivalent fixed base period index would be identical and hence give exactly the same results.

II Methods to be tested

Our approach is strictly descriptive and does not make any assumptions about market behavior while allowing all possible market interactions to be reduced into a joint distribution of all quality characteristics. We show how this same theoretical base may be used for apparently different quality correction methods.

A) Set-up
Without formal set-up we state that there is a time dependent joint distribution of price and quality characters of all goods considered. In this case we restrict ourselves to various digital camera models that are on the market at each period. A wider treatment of cross dependencies with other goods and services has been left out. Among other things, the unknown distribution fully describes the relation between the price of and all characters that influence the price. Unfortunately we can measure only a fraction of the characters and we have to leave most factors describing the camera markets out. Surely color, brad, appearance, production technology, market competition, regional differences, second hand markets, etc. affect the asking prices and transaction prices. These are, however, all included in the very complex multidimensional distribution.

We are interested in a particular conditional marginal distribution, or more precisely its' first central moment. This is the expected value of price conditioned on measurable quality characters $\boldsymbol{x}$ at certain time period. The conditioning has first been done over all other characters and we are left with a price function distribution. Next, we want to condition this on measurable characters and have the price as a function of these characters.

We describe the price function behind all types of indices simply by[234]

(1) $\left(h(P) \middle| x, t \right) = g^t\left(x\right) + \varepsilon^t,$

where the transformation function h is not generally known[235]. Because of nice properties of symmetry, summation over time, congruence with geometric means, and interpretation, we use natural logarithm as the transformation function for price in all hedonic regression models. Hence we base the price changes on log-changes. To simplify the notation we denote $\log P = p$. Empirical evidence also supports this choice at least in some cases, such as digital cameras here[236]. Now, we are interested in the systematic part of (1), i.e. the conditional expectation

(2) $E\left(\ln P \middle| x, t\right) = E\left(p \middle| x, t\right) = g^t\left(x\right).$

The function g gives the mean log-price for each combination of quality variables. At this point, nothing is assumed on the time specific functions g^t or the independent variables in x. Individual transformation functions (that may differ from identity or logarithm function) of measurable quality characteristics are applied to end up with vectors $x_i^t = \left(1, x_{1i}^t, \ldots, x_{Ki}^t\right)$ for each observation, each period. The K characteristics are treated as "measured in transformed form" and the question of model specification is not discussed further[237]. One should note that this representation still allows us to use flexible functional forms such as translog and quadratic models[238]. In all models, we assume the estimated function to be linear with respect to the parameters of the transformed original quality variables. Now, let function f be simply an estimate of the unknown relation g^t:

[234] One could also argue, as usually is done, that the real relation is an inverse of m which we transform this one just for sake of ease of estimation.
[235] This approach of estimated function follows Vartia and Koskimäki (2001).
[236] See e.g. Diewert (2002) and a summary treatment in Triplett (2002) or IMF (2002).
[237] For functional forms and economic approach to hedonic indices, see Rosen (1974), Diewert (2001) and Triplett (2002). Here the question of model selection and functional form is not discussed further.
[238] See Diewert (2001).

(3) $f^t(x) \equiv est\left[g^t\left(x\right)\right] = est\left[E\left(\log P \middle| x,t\right)\right]$.

For our purposes we always use the following linear (with respect to parameters) functional form:

(4) $f^t(x) = \hat{\beta}'^t\, x$,

where the *(K+1)* β–vector includes a constant term.

For the purposes of this study the estimation method used does not have to be OLS as long as it forces the sum of residuals to zero at each period[239]. Regardless of the estimation, we denote $\hat{p}^t_{x_i^{t-1}} = \hat{\beta}'^t\, x_i^{t-1} = f^t\left(x_i^{t-1}\right)$ as period *t* estimated price for period *t-1* observation *i,* and similarly $\hat{p}^{t-1}_{x_j^t} = \hat{\beta}'^{t-1}\, x_j^t = f^{t-1}\left(x_j^t\right)$ as period *t-1* estimated price for observation j in period *t.* Also, we denote the estimated log-price change as $\dot{p}$ for both $\hat{p}^t_{x^{t-1}} - p_i^{t-1} = \dot{p}_i$ and $p_i^t - \hat{p}^{t-1}_{x^t} = \dot{p}_i$. Alternatively, we could have used estimated prices on both periods[240].

We define the logarithmic hedonic price indices as weighted geometric means of estimated log-price changes of either base or reference period observations. The estimated log-prices are derived either using the reference period hedonic model for base period observations or base period model for the reference period observations. We have chosen to include also the weights in the table 1 below, although in the empirical part we will only use equally weighted means. The equally weighted geometric mean is also called Jevons index. We call these weighted indices hedonic log-Laspeyres (1) and log-Palgrave (2) respectively. By taking geometric mean of the two indices, we get the formula for hedonic Törnqvist index[241]. Other index number formulas based on log changes, such as Sato-Vartia or Walsh formulas are not investigated in this study.

Table 1 Hedonic logarithmic indices

Index	Formula	Weight	Estimated price
Log-Laspeyres (1)	$\log P_{t-1}^t = \sum_{i=1}^{N^{t-1}} w_i^{t-1}\left(\hat{p}^t_{x_i^{t-1}} - p_i^{t-1}\right)$	Period *t-1*	Period *t* using period *t-1* observations
Log-Palgrave (2)	$\log P_{t-1}^t = \sum_{i=1}^{N^t} w_i^t\left(p_i^t - \hat{p}^{t-1}_{x_i^t}\right)$	Period *t*	Period *t-1* using period *t* observations
"Törnqvist"	$P_{t-1}^t = \sqrt{\exp(1)\times\exp(2)}$	Both	Both

B) Decomposition of hedonic price index

This section provides an illustrative way of showing the quantitative effect of hedonic quality adjustment with geometric indices. The proof of formula (X) is given in appendix 2 which also

239 An example of trivial zero sum residual "estimation" of *g* is a relation *f* that gives the observed value for each observation *i*.

240 De Haan (2003) proposes the use of what he calls double imputation.

241 It is rather straightforward to show that the index refers to the Törnqvist formula only if the price relatives in (1) and (2) refer to the same (number of) observations. This is not true in the digital camera data. One cannot directly compute the true Törnqvist index without having the weights for each observation on both periods.

discusses the implications with patched model index. Some of these decompositions are presented in section seven together with regression results from the data.

It can be show that, regardless of the type of the hedonic model, hedonic log-Laspeyres index can be presented as

$$(15)\ P_{t-1}^{t}(La) = \exp\left[\sum_{i=1}^{N^{t-1}} \left(w_i^{t-1}\left(\hat{p}_i^t - p_i^{t-1}\right)\right)\right]$$
$$= \frac{G(P^t)}{G(P^{t-1})} \times \exp\left(\hat{\beta}'^{t}\left(\bar{x}^{t-1} - \bar{x}^t\right)\right) \times \exp\left(\operatorname{cov}\left(\frac{w}{\bar{w}}, \dot{p}\right)\right)$$
$$= \frac{G(P^t)}{G(P^{t-1})} \times \exp\left(\hat{\beta}'^{t}\left(\bar{x}^{t-1} - \bar{x}^t\right)\right) \times \exp\left[\operatorname{cov}\left(\frac{w}{\bar{w}}, \left(\hat{\beta}'^{t} - \hat{\beta}'^{t-1}\right)x^{t-1}\right)\right] \times \exp\left[-\operatorname{cov}\left(\frac{w}{\bar{w}}, e^{t-1}\right)\right].$$

This decomposition, though slightly differently, was first presented in Koev and Suopera[242]. It says that the quality adjusted log-Laspeyres index may be presented as a product of relative of geometric average prices at two periods, a separable quality correction term that depends on change in average quality, and a sample covariance term between mean adjusted weights and estimated pure log-price change. The last row of (5) breaks the covariance term into two parts; one that shows the effect of model differences at two periods and a term between the adjusted weight and period *t-1* estimation error. In case of a time indicator model, the first correlation term is zero, because the two models are same. If we used double imputation, also the second part of the covariance term would disappear. Then the hedonic log-Laspeyres index would be simply

$$P_{t-1}^{t}(La) = \frac{G(P^t)}{G(P^{t-1})} \times \exp\left(\hat{\beta}^{t}{}'\left(\bar{x}^{t-1} - \bar{x}^t\right)\right).$$

Also, in case of a time indicator model the coefficient estimate of β is not time specific and hedonic log-Laspeyres and hedonic Törnqvist indices are identical in formula. As said, the quality correction term may be separated into effects of each quality character and each effect may be presented separately:

$$(17)\ \hat{\beta}'\left(\bar{x}^{t-1} - \bar{x}^t\right) = \hat{\beta}_1\left(\bar{x}_1^{t-1} - \bar{x}_1^t\right) + \ldots + \hat{\beta}_K\left(\bar{x}_K^{t-1} - \bar{x}_K^t\right).$$

What is important here is that we can compare the quality correction factors of different model types to be introduced below. Hence, when comparing different quality adjustment methods we could, instead of the actual indices, compare the overall quality correction factors. If the factors are found to be sensitive to the choice of method (or regression model), the most accurate method should be used. The good new would be that, if the quality correction factor is fairly robust to the choice of method, index compiling agency may choose the method that can most easily be implemented in practice. Further, it is easy to see that with equally weighted indices, as elementary aggregate indices usually are, only changes in average quality matter.[243]

C) Methods

We separate six different types of hedonic models based on the model specification. Regardless of the model type, indices may always be calculated using the above formulas even if it may not be the

[242] Koev and Suoperä (2002).
[243] This decomposition may of course be extended to any log-based index formula and also mixed methods, such as patched model (see Pakes (2002). However, these are not discussed further in this study.

most efficient way. The different methods we are going to apply in the later sections are the following.

I) Grand unit-value
Quality adjustment is ignored and price index is simply the geometric average of price relatives. The model does not depend on the quality characters and reduces to a constant, but may still be estimated as a regression model:

$$(5) \quad f^t(x) = a^t .$$

The quality correction factor is simply 1.

II) Class unit-value
In this case some quality differences are taken into account. Class unit-value method is analogous to grouping the goods into subgroups and following the class averages. Using regression model we write

$$(6) \; f^t(x) = a^t + b_1^t D_1 + b_2^t D_2 + \cdots + b_C^t D_C = a^t + b_c^t = a_c^t$$

where D_c^t is a binary indicator variable for qualities in class c. The function *f* returns the mean price a_c^t in class c. The quality correction factor depends on chances in the distribution of observations into different classes. This can be seen from (x), where the change in average quality $\left(\bar{x}^{t-1} - \bar{x}^t\right)$ is now calculated from dummy variables representing the classes.

III) Matched model
As its name suggest, matched model follows the prices of same goods. This can be presented in terms of estimation function *f* returning the observed price for each period *t*. A possible notation of this is

$$(7) \quad f^t(x) = p_i^t \;\; \forall j \in \left(p \middle| x = x_i^t\right).$$

The quality correction factor is again just one, since there is no quality changes. It can be shown that for a patched model the quality correction factor is otherwise identical but also depends on the share of missing/new items and average estimation error between matched and unmatched observations.

IV) Time dummy pooled hedonic regression
This is the first 'real' hedonic model and maybe the most widely used hedonic specification. Quality characters are assumed to affect price similarly across time and the level transitions in time account for price chances. Using the first period as a reference the model may be written as

$$(8) \; f^t(x) = a + b'x + \beta_1 D^2 + \beta_2 D^3 + \cdots + \beta_{T-1} D^T = a + b'x + \beta_t = a^t + b'x .$$

The quality correction for all 'real' hedonic methods is given by (5).

V) Time dummy 2-period hedonic regression
The 2-period hedonic regression method is basically the same as the previous. The first difference is that instead of using all periods in the estimation, we now relax the temporary restrictions on the

quality coefficients over longer time period. Another difference is, although rarely considered, that we may change the quality characters included in the regression as well.

VI) Full hedonic imputation
This method lets the price-quality -relation to change freely from period to period. Leaving aside the matched model, all other methods are special cases of full hedonic imputation. The model is simply the same as in (4):

$$(9) \quad f^t(x) = a^t + b^{\prime t} x .$$

These are the methods, or model specifications that we are going to apply to the digital camera data. We expect that rapidly evolving product space and rapid technological change in quality would prove some of the methods clearly unsatisfactory and some clearly better. Before doing this we shortly discuss some general pros and cons of each method.

III Advantages and Disadvantages of the Methods

All six methods may be practical is some cases and some are clearly not. Some methods clearly depend on more information that has to collected.

I Grand unit-value
The grand unit-value index is a good method for homogenous items and it is used as elementary indices in some CPIs and PPIs. Especially, if the goods are truly homogenous and prices volatile the unit-value method may be appropriate. It escapes the difficulty of pricing point in time. However, in elementary aggregate case, the quality is more often controlled by the matched model approach, which would of course give the same result when calculated as a geometric mean.

II) Class unit-value
The Class unit-value method uses only categorical variables. In essence, by grouping the observations into subgroups we eliminate the quality differences between these subgroups, just as in analysis of variance. The quality adjusted price index may be constructed directly from class means implied by the classification (and cross classification).

III) Matched model
The matched model is more of an approach than a quality adjustment method. It is safe to same that majority of national CPIs and PPIs are based on the matched model way of thinking. The quality adjustment is seen as a separate problem only when we are missing the match. The advantage of collecting prices of same finely defined products from same stores are obvious. We can simultaneously control for some immeasurable quality characters without explicitly stating them.

The major drawback of the matched model approach is that the sample may become unrepresentative if not updated frequently enough. It incurs a sample selection bias. Matched model fails to account for new models that are introduced, as well as the old ones for which no matches can be found. If we don't just ignore the non-matches, some quality adjustment method is still needed.

Statistical agencies usually collect replacement items if the initial specification of a good are not found in consecutive periods. Then regression, or hedonic, methods are use to adjust the price of this new (or the missing) item.

IV) Time dummy pooled hedonic regression
This is the "classical" use of regression analysis with price indices and is what people often think when they perceive hedonics.

There are clearly some good properties that make time dummy models useful. First, the calculation is fairly simple and the index may also be constructed solely based on the estimated time indicator coefficients. Especially in research type studies that examine historical data, compared to updating an index series form time to time, Secondly, as Tripplett (2002) proposes, the regression coefficient in the reduced regression models are often relatively stable and time indicator models work fairly well.

However, there are drawbacks as well. One should always confirm that the coefficients indeed are stable. With high frequency indices and rapidly evolving commodity population sudden changes may go unnoticed.

The time indicator model does not have to be Using the formulas in table 1 including the weights, we can derive a "Törnqvist" index for time indicator model as well. It is needless to say that the resulting index is identical to log-Laspeyres and log-Palgrave indices.

V) Time dummy 2-period hedonic regression
Now instead of estimating just one regression model for 20 period data we estimate 19 models. The index could also be deduced directly from the time dummy indicators.

Time dummy pooled and 2-period pooled regression. The quality adjusted price index may be constructed directly from the estimates of time coefficients and, as mentioned before, the quality variables may be binary or continuous.

VI) Full hedonic imputation
We argue that this should be the basic starting point for all hedonic price index studies. From the standpoint of the resulting index all previous methods are special cases or simplifications of full hedonic imputation method. This, of course requires that the weight data is either derived from elsewhere or the data used for the regression model does not use e.g. permutations of "same items".

In addition to these basic methods, one may want to use combinations of them. E.g. it may not be feasible to estimate the whole regression separately each period because data restrictions. However, it may be imperative to let the coefficients of some central. An example in the digital camera context could be that the coefficient for picture accuracy (measured in megapixels) changes in time while the effect of brand stays constant in medium term. In the framework above this may be done simply by defining the estimation function accordingly. Although this was tried, the results are not reported in this article since no remarkable differences were observed. All methods give reasonable results if the assumptions are sufficiently fulfilled.

Hedonic methods that are used as structural models to "correct" the observed price for changes in some quality characters do not easily fall into any of these categories. However, it may also be written in terms of the above estimation function $\boldsymbol{f}$ so that for observation $\boldsymbol{i}$ it returns the observed price plus quality correction

$$f^{t}(x) = p_i^t + b'\left(x_i^t - x_i^t\right).$$

A variation of this method and used with missing observations is patched model (See Pakes, 2002), that avoid the dangers of structural correction. Patched model uses imputation for those observations that are missing, or where the quality has changed, for the period in question and matching prices for observations that can be found in both periods[244].

We don't want to promote the use of hedonics as in (12) and assign interpretations for the regression coefficients. We started from the complex joint distribution of all measurable and immeasurable characters and come down to, at best a simplified, set of characters and a reduced form of expectation. By this time, we cannot really claim that the individual coefficients themselves represent, if we may borrow from economic jargon, "shadow prices" without any knowledge of the true distribution. The price forecast, on the other hand, is just as good as the reduced form regression model.

IV Description of the Data Set

The quarterly price data were recorded from various sources, mainly from issues of Journal of Popular Photography on microfilm. For two last quarters of 2002, the price data were collected from the internet at *www.pricescan.com*. The model specific quality characteristics were compiled mainly from the website *dpreviw.com*.

Total of over 1300 prices of 288 different digital camera models were collected. Some of these prices were averaged over advertising retailers at the time of entry. Although the share of these multiple observations was not recorded, it accounts for approximately 10 to 15 per cent of the total sample. Regardless of the number of observations used for each recorded model price quote, they are treated as single observations.

I) The sample design

It was not feasible to use random sampling in the study because of scarce data sources. For the early years, all possible models with a price quote in the Journal were recorded. When the price collection was changed to internet, generally the lowest price was recorded and almost all available makes and models were included. There were no effort to follow same models, and market entries and exits occurred when new models were first advertised or they were no longer available. In other words, the sample was not designed to mimic any typical statistical agency approach.

After the data collection, we decided to exclude SLR[245] type digital cameras and also restrict the time sample to the years 1999 through 2002. The SLR-type digital cameras are used by professionals and they are much more expensive and have partly different properties and characteristics. The data from years 1996 and 1997 proved to be too limited.

The quality characteristics may have some variation between retailers. However, just one reference attribute is used throughout the lifespan of a model over all retailers. These differences could not be

[244] The term Patched model was used by Pakes (2001).
[245] SLR - single lens reflex.

observed in detail and are most likely limited to different memory cards included in cameras. For this reason, memory was excluded from some of the regression models.

The number of priced models (and hence price observations) follow the number of adds in the Journal and the number of different models each manufacturer makes. One might argue that the data is self-weighting in a sense that the more models a manufacturer has, the larger its share is in the data. No explicit data was available for the weights and all indices are calculated as equally weighted geometric means. In digital camera case, we assume that the number of different models advertised is a proxy for market share of that make. This view may be challenged, but no other data was available to support or contradict this hypothesis.

II) Data reservations

There are some issues in the data that might be a hindrance to the index:

1) *"Call for price"*. A number of times the Journal adds do not show the price directly, and they could only be obtained by calling the retailer. This practice is still in use and may be set by either the manufacturer or the retailer.
2) *Weight data*. There were no data on model nor manufacturer turnover or other data on relative importance of manufacturers.
3) *Unbalanced samples*. The price sample is unbalanced towards the last year and the last two quarters when the price collection changed fundamentally.
4) *Limited data on the early years.*

These concerns were treated by the following practices, which are commonly used. However, their effects should always be estimated.

1) *"Call for price"*. These prices were not collected at all. Our assumption is that the "Call for price" is more widely used with new introductory models than older models. Although the basis for reasoning is not relevant in our descriptive indices, one may argue that models entering the market have higher mark-ups while models exiting are being sold out with smaller or even negative markups. Since we are not referring to marginal cost –pricing, as we are only interested in the actual prices paid, we assume that there are no differences in the price determining processes (or marginal distributions) between the advertised and non-advertised prices.
2) *Weight data*. As already noted, since we had no data on quantities sold by model or by manufacturer, we assumed that the number of models advertised provides a proxy for the manufacturer weights. Thus, each model has an equal weight. Out of all observations Sony counted for most (18%) and then Olympus (15%), Kodak (13%), Fuji (12%), Canon (11%) and Nikon (8%) of total observations. The rest 13 manufacturers account for the rest 24 % of observations. This self-weighting seems reasonable – especially since there were not great differences in pricing between manufacturers.[246] However, it would be very interesting to see how model specific weights would have affected the series.
3) *Unbalanced samples*. There is a clear change in the number of observations per quarter when the data source was changed to the internet after the second quarter of 2002. There may be a systematic reduction in prices due to this change, since the data selected from Pricescan.com usually refer to "Best price" which is the lowest advertised price within a selection of online retailers. These prices usually do not include shipping and there may be a tendency for the low price retailers to charge more for the shipping than others. However, this is clearly not

[246] As tested with classification models.

the case every time and the same may also be true for the Journal adds as well. The degree of possible bias from this has not been quantified. There were total of 1052 price quota for the years 1998 – 2002 and the distribution of observations is presented in table 3.1.

Table 3.1 The distribution of price observations

Quarter / Year	1998	1999	2000	2001	2002
Q1	19	41	79	75	50
Q2	20	34	44	29	94
Q3	27	54	63	83	155
Q4	37	31	55	41	124
Total	103	160	241	228	423

The change in the data collection is likely to have an effect also on the index. The prices collected from the internet are usually the lowest prices for each model and not averaged over advertised prices, as is the case in the price data from the Journal adds. By collecting overlapping prices, one could estimate the magnitude of a shift change but this exercise was not carried out. Another feature of internet purchases is the shipping and handling fees. Although not included in the prices, it could be argued that in some cases a part of the actual price is actually charged as shipping and handling, which is rather evident with regard to some special offers of other consumer goods. However, a quick sample did not confirm the negative relationship between price of the good and the handling fees for same models.

4) *Limited data on the early years.* There are additional 21 price observations for 1996 and 38 for 1997 that have not been used in the analysis, because these data include missing information on the characteristics. With a distinctive model or other method the series could be extended a few years back, but for this study this exercise was not carried out. Also, if the time period were further reduced to, say the last three years, differences between quality adjusted and unadjusted price indices would not appear so large[247].

These questions, as important as they may be, are not addressed further in this study. Our main purpose was to compare different methods of controlling quality and the index series they produce. Some of the questions could have been solved with some complementary data or additional data collection.

III) Quality characters

The regression models to be laid out in section 5 need input variables, namely regressors. Potentially good explanatory variables for hedonic regression would be such that have variability in time and relatively large spread within each time period. Additionally, they should be correlated with the price and not heavily with each other. In figure 4.2, there are some potential variables to be used in the regression models in form of a simple average relative change thus showing only the variability in time.

[247] This will become apparent later.

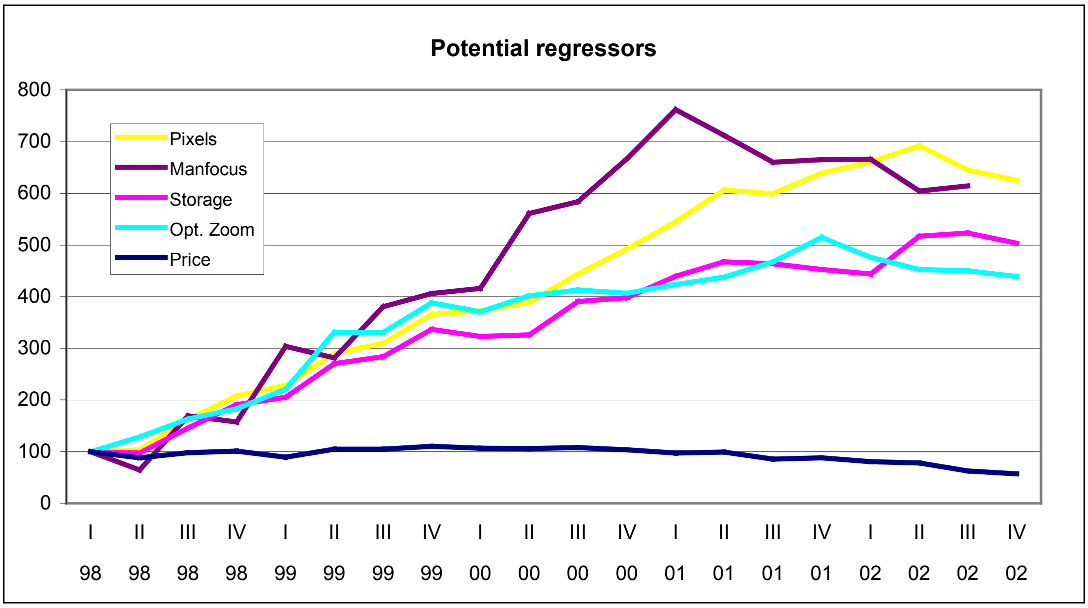


Figure 4.2 Some average quality variables as indices

Changes in the average quality characteristics are large in time compared to the evolution of average prices. Of these variables, the manual focus is a binary variable and the index presentation should be interpreted as the evolution of the share of digital cameras having the feature in question. When used in average form e.g. in imputation, all binary variables should be interpreted similarly. All available variables in the data set are presented in table 2 in Appendix 1. The most useful are the ones indicating a sharpness of the picture, a memory capacity, an optical zoom ratio, and manual focus, an external flash and movie options.

V Results

As a group digital cameras (compact or ultra compact models) is a rather homogenous set of product varieties (compared to some other transactions in the economy). For many high technology goods, the average price does not seem to change very much but at the same time average (technological) quality characteristics change considerably. This is true for digital cameras too.

I) Grand unit-value

The price for a typical new digital camera model often starts with a stable introductory price (may be set by the manufacturer) and then the dispersion of offer prices becomes larger in time. Often the highest asking price stays the same (or decreases moderately) while the lowest price declines sharply. In the data, there is just one price for a model at one time, but additional sales information would be interesting to obtain.[248]

Without controlling for quality changes the average price fluctuates around $500 until early 2001 and then drops to under $350 during the last year. The grand unit-value index refers to the geometric means. Changes in quarterly gross unit prices are presented in figure 4.1.

[248] See e.g. www.pricescan.com.This property brings some additional problems for the price index that were not accounted for in section 3. The data collected does not allow to take into account this behavior.

Figure 4.1 Unit value index series (1998 = 100)

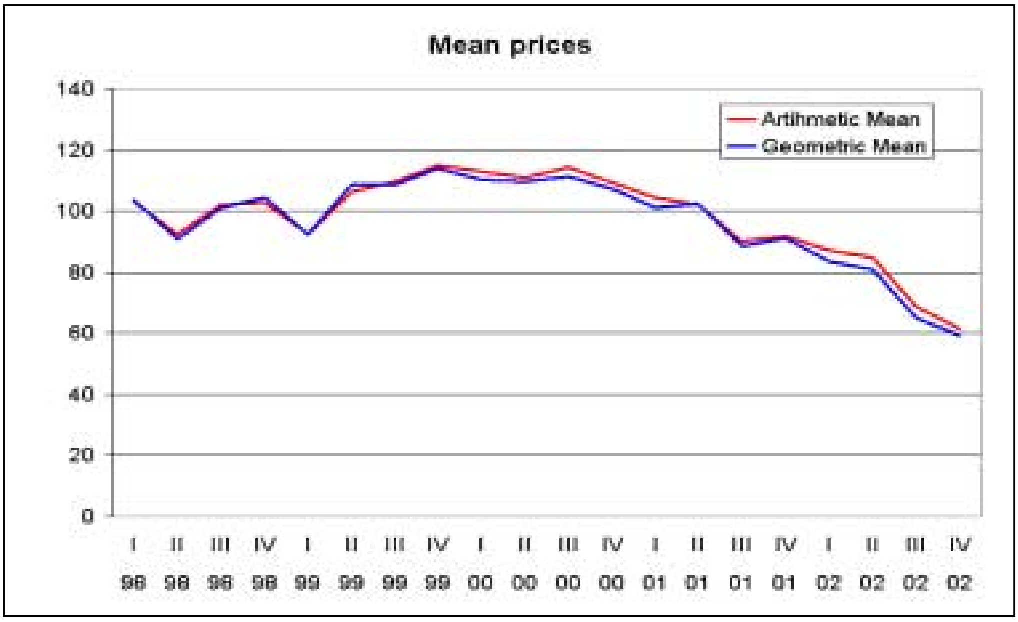


Figure 4.1 confirms the intuition we had on digital cameras. The prices of ‘a good camera’ have been around \$500 for some years and just during the last few years the prices have begun to drop, though less than the picture would indicate[249]. Of course, these prices do not count for the changes in performance and other characteristics of the equipment. Equally importantly, these average prices do not follow the prices of same camera models. The drop during the last few quarters may be partly explained with the change in data collection – both because the recorded price refers to the lowest price and also because there may be more ‘low end’ models in the data set.

In terms of the model in section 2, calculating unadjusted geometric means is equal to “estimating” a hedonic regression model

$$(13') \quad f^t(x) = \hat{\beta}^{\prime t} x^t = \hat{\alpha}^t .$$

This can be estimated from pooled data consisting all quarters and the resulting model fit measured in adjusted R^2 is little over 20% and most of individual t-test statistics for time indicator coefficients suggest that they do not deviate from *0*[250]. We call this a trivial model. It is actually used by some statistical agencies in imputing missing observations, though not usually explicitly in this form.

II) Class unit-value

The usual practice statistical agencies use in tackling the changes in quality comes as a by-product of the sampling or selection process. Just follow the prices of same goods in time! In case of digital cameras, the ‘same’ would mean the same store and same model of the same make. No measurable quality change to control for – by definition. This could be interpreted as using a tight classification where each model forms one class. The index is based on the class average prices. For the sake of illustration, we classify (ex post) the cameras to similar (homogenous) groups and calculate the group mean price changes. This classification method is actually often used by statistical agencies for missing observations or replacements and may be a good method in connection with some products.

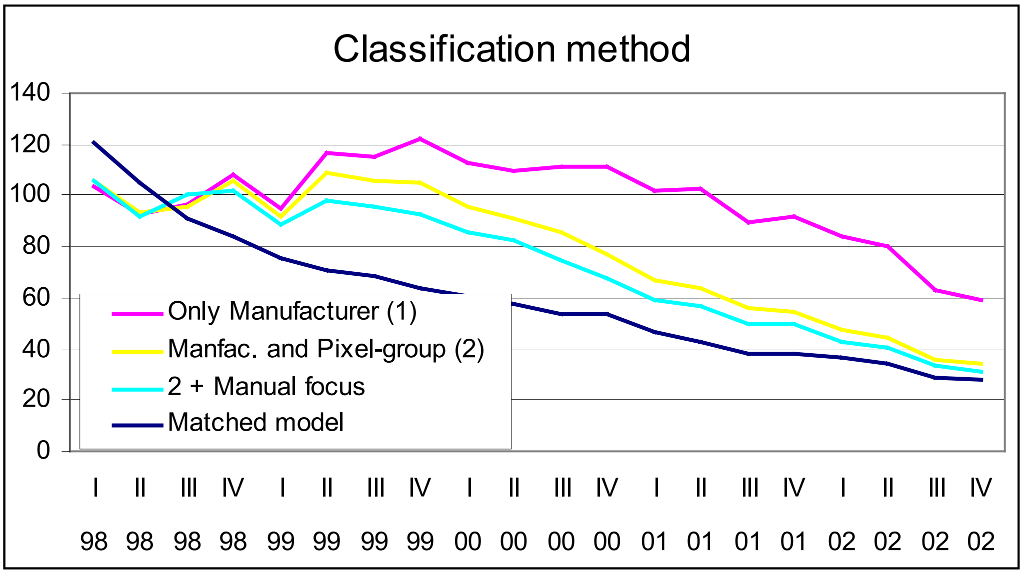
[249] We believe the same goes with some other high-tech equipment, e.g. PC’s. For some years now you’ve got a ‘good computer’ – not the very state of the art, but the next best thing – for some \$1500.
[250] Actually t-tests for H0-assumption of all individual α:s being zero would be rejected at 5% confidence level for only last two quarters if the first quarter is used as a reference.

The classification method is based on classifying the cameras according to some rule – most likely by their characteristics – and calculating the class means. The matched model index above is an extreme case of this method. Each model is classified as its own group and 'empty' classes appear every time when the model is not found in the next period.

A more realistic case would be to use, for example, the manufacturer as a classification rule. With the classification, one hopes to remove as much as possible of the within class price variation at each time. This method is typically used in repeated measurement experiments in natural sciences and called analysis of variance. It is usually not called that way in the price index context and usually none of the available test statistics from the analysis are either estimated or presented.

The first classification model in figure 5.2 is based on the manufacturer. The index may be calculated from the changes in make-specific average prices[251]. The second model adds pixel group – a variable, which classifies the camera models into five categories according to the available accuracy of the picture (less that 1 megapixel, 2 Mp, 3Mp and over 4 Mp cameras). The third model further classifies the data to models with or without a manual focus option (autofocus is the norm). As with unadjusted averages, we do not actually calculate the class means but instead use regression models (without cross effects).[252]

FIGURE 5.2 THE CLASSIFICATION METHOD



Compared to the matched model index, these indices have more volatility around the trend in the first two years. Adding the classifying factors clearly smoothens the series, but practical usability suffers because the number of class means to be calculated grows exponentially and empty classes start to appear.

The classification method may be presented in the notation of section II as follows:

$$(14)\ f^t(x) = \hat{\beta}'^t x^t = \alpha + \beta_1 D_{MANUF} + \beta_2 D_{MF} + \cdots + \sum_{t=2}^{T} \gamma^t D_t,$$

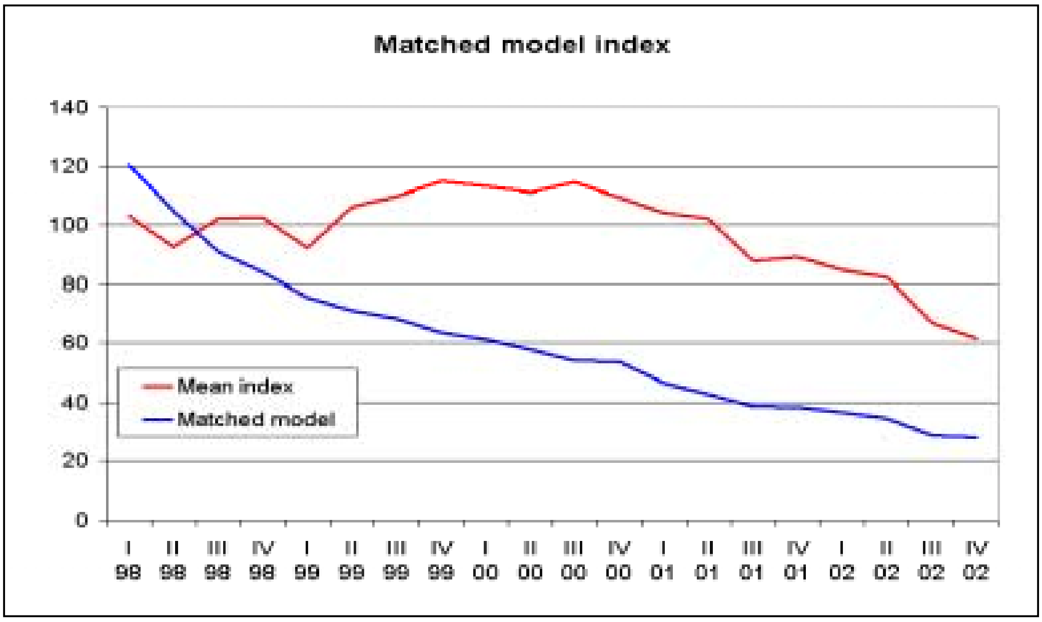
[251] The actual calculation is based on a time dummy regression model with indicator variables for each manufacturer.
[252] So we are actually not cross-classifying the models, but only using the 'main effects'.

where Ds are binary variables for classifying variables and time indicator variables. When estimated from pooled data the R-squares for the three models above are 33, 62 and 69% respectively. As already mentioned, the index series may be constructed directly from the time indicator coefficients. However, the same resulting series is also reached by estimating the price for each observations and using the index number formulas presented in table 1.1.

III) Matched model
Before using the classification method, we first calculate a digital camera price index based on a matched model method. As mentioned above in section 3.1, the sampling frame was not meant to be used for calculation of a statistical agency –type matched model index. In a sense, this sampling could be described as "quarterly re-sampling"[253]. Since we did not initially plan to calculate the matched model index at all, it is provided only as an example and methodological criticism should not focus on inadequacy of this matched model method. Also, no patched models were compiled from the data.

There were observations from at least two quarters for almost all of the 288 models in the data set. However, the turnover of camera models was rather fast. The number of models for which price were found over more than 4 quarters' period was 167 and over 6 quarters' just 52. In traditional statistical agency practice, this would have meant a large number of replacement models to be found and an alternative method to account for those models at times of no price observation. Our matched model index does not include estimates for the missing models, either for the ones entering or exiting. The observed price change over more than one quarter is divided by the number of quarters and addressed only to the first quarter. With these reservations, the (log mean) index series are presented in figure 4.1 together with an unadjusted average price index. As can be seen, especially in the first six to eight quarters, the two methods differ considerably, and the matched model index is much smoother in decline. We will get back to some interpretation in section 6.

Figure 5.1 The matched model index and simple average (1998 = 100)



[253] It would not be true to claim that the samples were truly independent from one another since we used the same magazine having mostly the same advertisers over time. However, the notion of independence should not be too far from true and we would expect it to have only minor effect for the price index.

As already noted in section II, all models use natural logarithmic of price for the dependent variable and either logarithmic or identity transformation functions for the explanatory variables. In the first sub-section we introduce the index formula decompositions. I the next section we estimate models restricting the coefficients of all quality characteristics to be constant over time, while in 5.3 these restrictions are relaxed.

IV) Time dummy pooled hedonic regression

In this section all period t models are estimated as

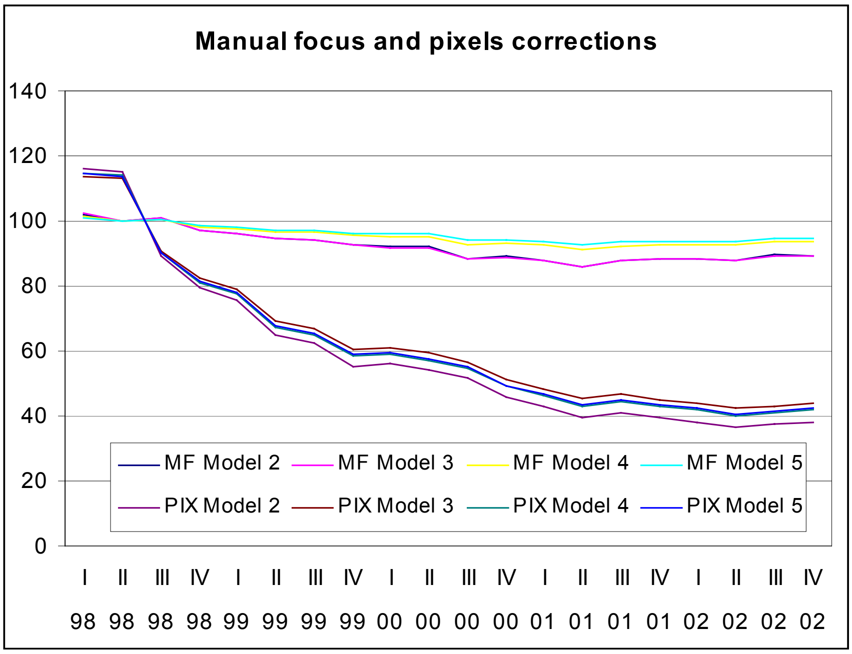
$$(18)\quad \ln\left(\hat{p}_i^t\right) = \hat{\beta}^t x_i^t = \hat{\alpha} + \hat{\beta}_1 x_{i1}^t + \cdots + \hat{\beta}_K x_{iK}^t + \sum\nolimits_{t=1}^{T} \hat{\delta}^t D_t \ ,$$

where an indicator variable D_t gets value 1 at period t and 0 otherwise. We will use four slightly different models to illustrate how the model selection affects the quality correction and the index. The models are:

Model 1: manual focus + ln(megapixels) [R-square 66%]
Model 2: manual focus + ln(megapixels) + ln(megabytes) [75%]
Model 3: manual focus + ln(megapixels) + optical zoom [77%]
Model 4: manual focus + ln(megapixels) + optical zoom + ln(megabytes) + external flash [80%]

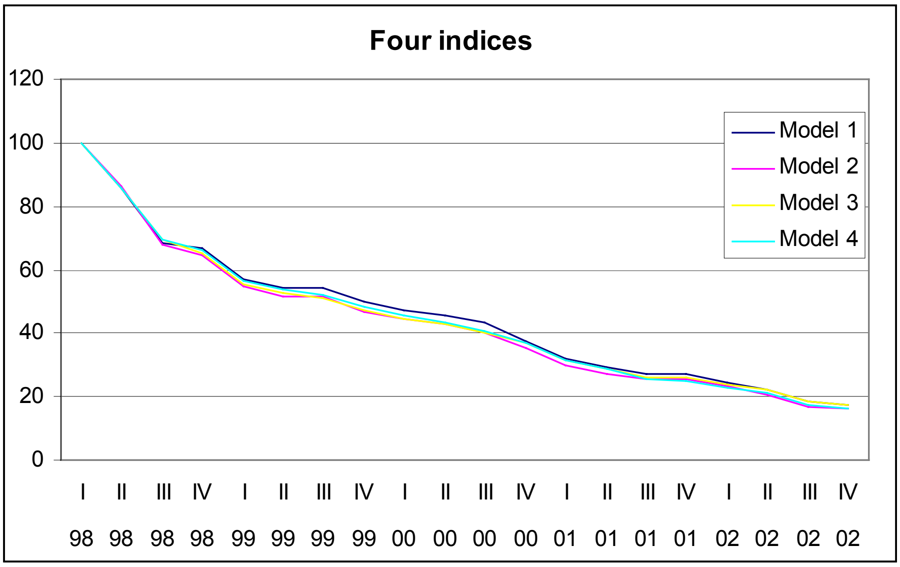
Adding more quality characteristics increase the overall model fit somewhat, and the differences with quality correction factors from each model become even smaller. A summary table for the coefficients and quality correction terms of the below models are presented in appendix 4. In figure 6.1 quality correction terms for two characteristics, manual focus and pixels, are presented.

Figure 6.1 Quality correction factors for four models



As one can see there is some variation between the four models when the OLS adjusts the hyperplane in price-quality coordination. The feature of forecast model is that by adding explanatory variables into the model the individual coefficients, and quality correction factors, adjust so that best overall fit is achieved. This means that the individual quality corrections factors contribute a part of their value to the new variable depending on the amount of multicollinearity it has with the variables already in the model. However, when taken together with all quality characteristics in the model, the total quality correction factor may have very little 'dispersion' between models, as in figure 6.2.

Figure 6.2 Quality adjusted log-Laspeyres indices for the four models



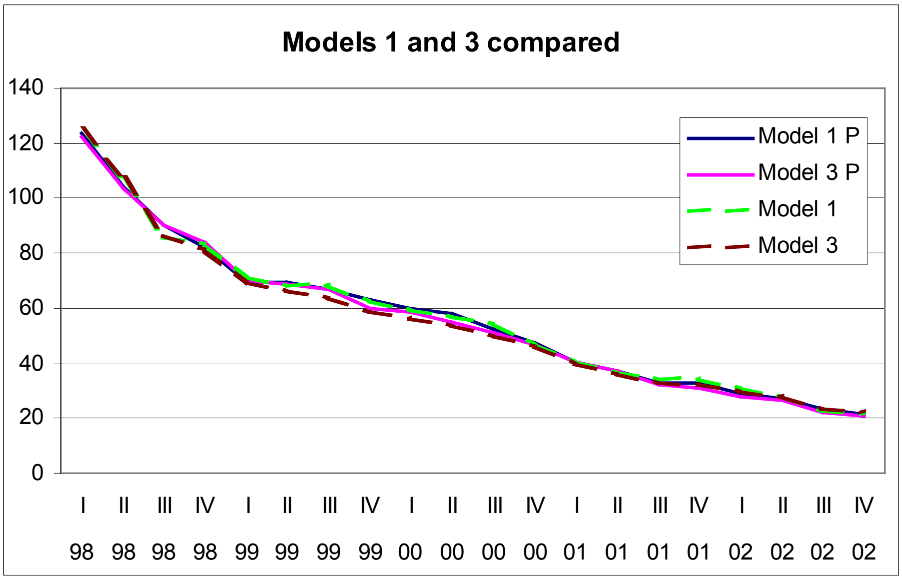
The overall picture of the quality adjusted price index for 1999 – 2002 does not change when adding new quality variables into the hedonic regression. Also, for practical reasons it may be feasible to collect even high frequency data on just few quality characteristics together with an existing price collection, e.g. the CPI.

V) Time dummy 2-period hedonic regression

In case of rapid quality change, the assumption that quality – price relation stays the same except from the constant term over a long period of time should be questioned. We will use two different methods to allow more flexibility in the models. The first one is to apply the time indicator model to data that pools data together only two consecutive periods, and estimate 19 independent time indicator models (for all pairs). The second is to estimate separate models for each 20 quarters and impute the matching prices as suggested by the index formulas. With the latter we calculate different log-Laspeyres and "log-Palgrave" indices using the model from period t and t-1, respectively and present the hedonic Törnqvist price index. We call these models pairwise pooled and full hedonic models.

As the results will show, again there are no large changes in the quality adjusted indices and thus we will use just two models, Model 1 and Model 3 from the previous section.
Now, the model R2:s vary between 60 and 85%. See a summary table of estimation results in Appendix 4. The resulting index series for pairwise pooled indices are presented in figure 6.3 together with ones from the previous section. The P refers to pairwise model, and as one can see, the two quality correction magnitudes are very similar with the completely pooled data models.

Figure 6.3 Pooled estimation models



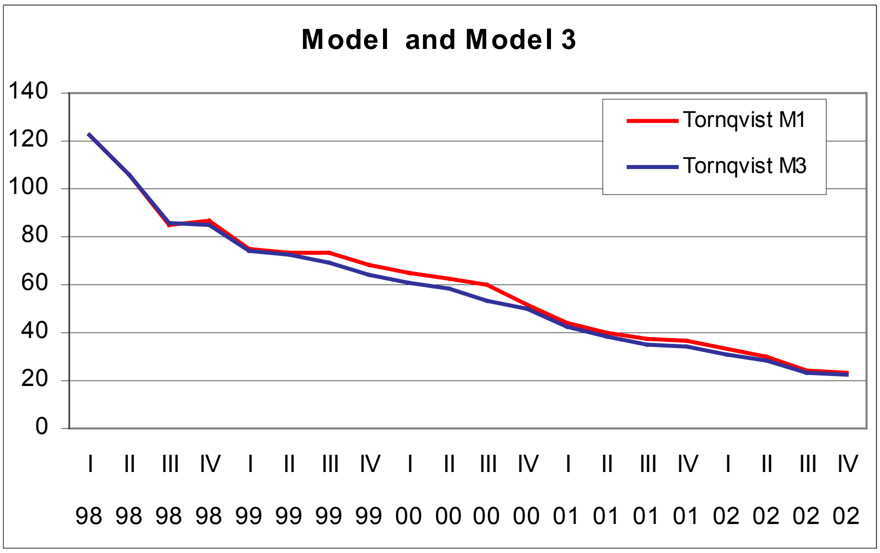
In this case individual observations have much more effect for the coefficients, especially in the early years. Consequently, the quality correction factors for individual characteristics do vary little more, but the total quality correction factors are not affected as much, as expected. One could modify this method by adding the number of consecutive periods to the estimation, which would have a further smoothing effect but still gradually take into account possible changes in quality – price relation.

VI) Full hedonic imputation

The second method could be called a true full hedonic method. Estimated models stay the same but instead of calculating just one index we will use all the data and present the resulting quality adjusted price indices as a Tornqvist index[254]. This would theoretically be the most comparable with an index based on time indicator model[255]. See result in the summary table in Appendix 4. The two indices are presented in figure 6.4.

[254] Difference of chained Laspeyres and Palgrave are at most 7 index points and average to very close to 0.
[255] Since both use the data from the two periods to estimate the model(s). See appendix 1.

Figure 6.4 Full hedonic models



While it may be difficult to see real difference in the quality adjusted price indices, there are some differences in the quality correction factors. The difference between the two model pairwise quality correction factors is some 8% at maximum. With the pooled data the difference is at most 6%. With individually estimated Fisher quality correction factors between the two model differ again at most some 8%.

VII) Summary
Tähän yhteenvetotaulu

VI Conclusions

As shown above, regardless of different methods and hedonic models used to adjust for quality changes in digital camera price index, all reasonable models produce indices that come very close to each other. Since we apply a forecast model we prefer a simpler model if no substantial benefit is achieved from adding more variables into the hedonic model. If one has to choose between a simple model and more complicated one we think the simpler is better.

As a starting point we should expect and allow the models differ between periods. Only if there is no strong evidence against changing coefficients, we may use time restricted models. If data allows the use of pooled regressions, hedonic models seem to be rather robust in choice of estimation and of model specification. The advantages from using somehow pooled data are that it gives more stable regression coefficients and 'smoothens' the quality adjusted index series. Especially with high frequency indices it may also make the hedonics more feasible since it demands less data. This is again not a bad thing.

Since relatively simple models may work well enough, hence large scale characteristics collection may not be necessary, and may not be any greater restraint for statistical agencies than a matched model approach.

As long as the matched model index does not produce outside the sample bias it works very well. However, if the distribution of characteristics changes in time, as it does with high technology products, frequent sampling is needed and quality changes in mismatches must be dealt somehow.

Typical statistical agency procedures may not be suitable for simultaneously dealing with quality change and sampling. We argue that hedonic approach is a good and often feasible way to produce indices so that sampling may be separated from the quality adjustment process.

The main findings indicate that, in an aggregate context, such as price index, relatively simple hedonic models may be sufficient for accurate quality controlling even in high technology products. Further, if compared with a matched model framework, the collection of characteristics data for hedonics may not need to exceed the precision already needed to "make the match". This suggests that it may be feasible to use hedonic indices even in high frequency index compilation.

As Pakes (2004), we see no real obstacles for using some form of hedonic approach instead of more fixed matched model approach. To validate this, we claim that additional cross sectional explanatory power from a set of added quality characteristics in hedonic models have only marginal longitudinal effects in the index series.

References

de Haan, J. (2003), Direct and Indirect Dummy Approaches to Hedonic Price Measurement, Revised version of paper presented at International Working Group on Price Indices, Paris

Diewert, W. E. (2001), Hedonic Regressions: A Consumer Theory Approach, Discussion Paper No.: 01-12, University of British Columbia, Vancouver

Diewert, W. E. (2002), Hedonic Producer Price Indexes and Quality Adjustment, Discussion Paper No.: 02-14, University of British Columbia, Vancouver

Heravi, S. and Silver, M. (2002), On the Stability of Hedonic Coefficients and their Implications for Quality-Adjusted Price Change Measurement, Cambridge, Mass: NBER Summer Institute 2002.

Hulten C. R. (2002): Price Hedonics: A Critical Review. NBER Summer Institute 2002, Cambridge, Mass.

Feenstra, R.C. (1995), Exact Hedonic Price Indexes, Review of Economics and Statistics, LXXVII, 634-54.

Gordon, R.J. (1990), The Measurement of Durable Goods Prices, Chicago: University of Chicago Press

IMF (2002): PPI Manual

Koev, E. and Suoperä, A. (2002), Omakotitalojen ja omakotitalotonttien hintaindeksit 19985=100, Statistics Finland report.

Koskimäki, T. and Vartia, Y. (2001),

Vartia, Y.O. and Vartia, P.L.I. (1984), Descriptive Index Number Theory and the Bank of Finland Currency Index, Scandinavian Journal of Economics 86 (3), 352-64.

Pakes, A. (2002): A reconsideration of Hedonic Price Indexes with an Application to PC's, Harward University.

Rosen, S. (1974): Hedonic Prices and Implicit Markets: Product Differentation in Pure Competition

Silver, M. and Heravi, S. (2001), Scanner Data and the Measurement of Inflation

Triplett, J.E. (2002), Handbook on Quality Adjustment of Price Indexes for Information and Communication Technology Products, OECD Directorate for Science, Technology and Industry, Draft, OECD Paris

Appendix 1. Decomposition of hedonic geometric indices

It was proposed that hedonic log-Laspeyres index may be decomposed as:

$$(A1)\quad P_{t-1}^{t}(La) = \exp\left[\sum_{i=1}^{N^{t-1}}\left(w_i^{t-1}\left(\hat{p}^{t}_{x_i^{t-1}} - p_i^{t-1}\right)\right)\right]$$

$$= \frac{G(P^t)}{G(P^{t-1})}\times\exp\left(\hat{\beta}'^{t}\left(\overline{x}^{t-1}-\overline{x}^{t}\right)\right)\times\exp\left(\operatorname{cov}\left(\frac{w}{\overline{w}},\dot{p}\right)\right).$$

To see this, we develop the basic hedonic log-Laspeyres formula by adding and subtracting estimated equally weighted pure price change

$$(A2)\quad P_{t-1}^{t} = \exp\left[\sum_{i=1}^{n}\left(w_i^{t-1}\left(\hat{p}_i^{xt} - p_i^{t-1}\right)\right)\right]$$

$$= \exp\left[\sum_{i=1}^{n}\left(\left(w_i^{t-1}-\frac{1}{N^{t-1}}\right)\left[\hat{p}^{t}_{x_i^{t-1}} - p_i^{t-1}\right]\right)+\frac{1}{N^{t-1}}\sum_{i=1}^{n}\left(\hat{p}^{t}_{x_i^{t-1}} - p_i^{t-1}\right)\right]$$

$$= \exp\left[\sum_{i=1}^{n}\left(\left(w_i^{t-1}-\frac{1}{N^{t-1}}\right)\left[\hat{p}^{t}_{x_i^{t-1}} - p_i^{t-1}\right]\right)\right]\times\exp\left(\frac{1}{N^{t-1}}\sum_{i=1}^{n}\left(\hat{p}^{t}_{x_i^{t-1}} - p_i^{t-1}\right)\right).$$

Now manipulating the second part by again adding and subtracting period *t* average price it may be written as:

$$(A3)\quad \exp\left[\frac{1}{N^{t-1}}\left(\sum_{i=1}^{N^{t-1}}\hat{p}^{t}_{x_i^{t-1}} - \sum_{i=1}^{N^{t-1}} p_i^{t-1}\right)\right]$$

$$= \exp\left(\hat{\beta}'^{t}\,\overline{x}^{t-1} - \overline{p^{t-1}} + \left(\overline{p^{t}} - \overline{p^{t}}\right)\right)$$

$$= \exp\left(\overline{p^{t}} - \overline{p^{t-1}}\right)\times\exp\left(\hat{\beta}'^{t}\,\overline{x}^{t-1} - \hat{\beta}'^{t}\,\overline{x}^{t}\right)$$

$$= \frac{G\left(P^{t}\right)}{G\left(P^{t-1}\right)}\times\exp\left(\hat{\beta}'^{t}\left(\overline{x}^{t-1}-\overline{x}^{t}\right)\right).$$

The first term is the relative of geometric means of the two periods' prices. It is an unadjusted or unit price index from period *t-1* to *t*. The second term is a multiplicative quality correction term that may be further factored into each characteristic. After estimating the regression coefficients, this equally weighted decomposition may easily be used for index calculation, since the quality correction term only depends on the average quality change.

The second term of (A1) may be written as

$$(A4)\quad \exp\left[\sum_{i=1}^{N^{t-1}}\left(w_i^{t-1}-\frac{1}{N^{t-1}}\right)\left[\hat{p}_{x_i^{t-1}}^{t}-p_i^{t-1}\right]\right]$$
$$=\exp\left[\frac{1}{N^{t-1}}\sum_{i=1}^{N^{t-1}}\left(\frac{w_i^{t-1}}{\overline{w}^{t-1}}-1\right)\left[\hat{p}_{x_i^{t-1}}^{t}-p_i^{t-1}\right]\right]$$
$$=\exp\left[\operatorname{cov}\left(\frac{w^{t-1}}{\overline{w}^{t-1}},\left[\hat{p}_{x_i^{t-1}}^{t}-p^{t-1}\right]\right)\right]$$
$$=\exp\left[\operatorname{cov}\left(\frac{w}{\overline{w}},\dot{p}\right)\right]$$

Also, to emphasize the hedonic model, we could further decompose the covariance term into *systematic and random* parts. Since the average weight depends on the number of observations, we may also write $\operatorname{cov}\left(w/\overline{w},\dot{p}\right)=\operatorname{cov}\left(N^{t-1}w,\dot{p}\right)$. To see the effect of selecting the model type this covariance term may further be written as

$$(A5)\quad \exp\left[\operatorname{cov}\left(\frac{w^{t-1}}{\overline{w}^{t-1}},\left[\hat{p}_{x_i^{t-1}}^{t}-p^{t-1}\right]\right)\right]$$
$$=\exp\left[\operatorname{cov}\left(\frac{w}{\overline{w}},\left[\hat{p}_{x_i^{t-1}}^{t}-\hat{p}_i^{t-1}-e_i^{t-1}\right]\right)\right]$$
$$=\exp\left[\operatorname{cov}\left(\frac{w}{\overline{w}},\hat{p}_{x_i^{t-1}}^{t}-\hat{p}_i^{t-1}\right)+\operatorname{cov}\left(\frac{w}{\overline{w}},-e_i^{t-1}\right)\right]$$
$$=\exp\left[\operatorname{cov}\left(\frac{w}{\overline{w}},\left(\hat{\beta}'^{t}-\hat{\beta}'^{t-1}\right)x^{t-1}\right)+\operatorname{cov}\left(\frac{w}{\overline{w}},-e_i^{t-1}\right)\right].$$

Now, it’s easy to see that for any time indicator model, for which regression coefficients stay constant, the covariance is between the weights and forecast error. Errors of course depend on how many periods are used in the estimation of time indicator model.

To derive a symmetric hedonic index that makes use of both period weights and regression models we start with repricing period *t* observations using period *t-1*model. It is easy to show that this Palgrave type formula is almost the same:

$$(A6)\ P_{t-1}^{t}(Pa)=\exp\left[\sum_{i=1}^{N^t}\left(w_i^t\left(p_i^t-\hat{p}_{x_i^t}^{t-1}\right)\right)\right]=\cdots$$
$$=\frac{G\left(P^t\right)}{G\left(P^{t-1}\right)}\times\exp\left(\hat{\beta}^{t-1}\left(\overline{x}^{t-1}-\overline{x}^{t}\right)\right)\times\exp\left[\sum_{i=1}^{N^t}\frac{1}{N^t}\left(w_i^t-\frac{1}{N^t}\right)\left[p_i^t-\hat{p}_{x_i^t}^{t-1}\right]\right]$$
$$=\frac{G\left(P^t\right)}{G\left(P^{t-1}\right)}\times\exp\left(\hat{\beta}^{t-1}\left(\overline{x}^{t-1}-\overline{x}^{t}\right)\right)\times\exp\left(\operatorname{cov}\left(\frac{w}{\overline{w}},\dot{p}\right)\right).$$

Finally, the hedonic Törnqvist index may be obtained as a geometric average of the two. The traditional Törnqvist index formula uses arithmetic mean of weights as

$$(A7)\ P_{t-1}^{t}(T\ddot{o}) = \exp\left(\sum_{i=1}^{N}\left(w_i^t + w_i^{t-1}\right)/2\left(p_i^t - p_i^{t-1}\right)\right).$$

BUT SINCE WE MAY HAVE A DIFFERENT NUMBER OF OBSERVATIONS IN THE TWO PERIODS WE DEFINE THE HEDONIC TÖRNQVIST INDEX AS GEOMETRIC MEAN OF HEDONIC LOG-LASPEYRES (A1) AND THE CURRENT PERIOD WEIGHTED "PALGRAVE-TYPE" INDEX (A6). USING NOTATION $\bar{\hat{\beta}}$ FOR AVERAGE OF THE TWO PERIOD ESTIMATED REGRESSION COEFFICIENTS IT IS RATHER STRAIGHTFORWARD TO SHOW THAT:

$$(A8)\ P_{t-1}^{t}(T\ddot{o}) = \sqrt{P_{t-1}^{t}(La)\times P_{t-1}^{t}(Pa)} = \frac{G\left(P^t\right)}{G\left(P^{t-1}\right)}$$

$$= \frac{G\left(P^t\right)}{G\left(P^{t-1}\right)}\times\exp\left(\bar{\hat{\beta}}'\left(\bar{x}^{t-1} - \bar{x}^{t}\right)\right)\times\exp\frac{1}{2}\left(\operatorname{cov}\left(\frac{w^{t-1}}{\bar{w}^{t-1}}, \hat{p}_{x^{t-1}}^{t} - p^{t-1}\right) + \operatorname{cov}\left(\frac{w^t}{\bar{w}^t}, p^t - \hat{p}_{x^t}^{t-1}\right)\right),$$

where the last term could be simplified - if the number of observations would stay the same – as

$$(A9)\ \frac{1}{2}\left(\operatorname{cov}\left(\frac{w^{t-1}}{\bar{w}^{t-1}}, \hat{p}^t - \hat{p}^{t-1}\right) + \operatorname{cov}\left(\frac{w^t}{\bar{w}^t}, \hat{p}^t - \hat{p}^{t-1}\right)\right) + \frac{1}{2}\left(\operatorname{cov}\left(\frac{w^{t-1}}{\bar{w}^{t-1}}, -e^{t-1}\right) + \operatorname{cov}\left(\frac{w^t}{\bar{w}^t}, e^t\right)\right)$$

$$= \operatorname{cov}\left(\frac{1}{2}\frac{w^{t-1}}{\bar{w}^{t-1}} + \frac{1}{2}\frac{w^{t-1}}{\bar{w}^{t-1}}, \hat{p}^t - \hat{p}^{t-1}\right) + \frac{1}{2}\left(\operatorname{cov}\left(\frac{w^t}{\bar{w}^t}, e^t\right) - \operatorname{cov}\left(\frac{w^{t-1}}{\bar{w}^{t-1}}, e^{t-1}\right)\right)$$

$$= \operatorname{cov}\left(w_T, \dot{p}\right) + \frac{1}{2}\left(\operatorname{cov}\left(\frac{w^t}{\bar{w}^t}, e^t\right) - \operatorname{cov}\left(\frac{w^{t-1}}{\bar{w}^{t-1}}, e^{t-1}\right)\right).$$

In (A9) the weights are now Törnqvist weights and also the regression coefficients used in the quality adjustment term follow the Törnqvist, in a sense that they are arithmetic means. Actually, any index formula based on log-change may be decomposed in the above way, just the covariance terms associated with the weighting scheme change.

Appendix 2. Average prices and the classification method

TABLE A1.1 AVERAGE PRICES

Year	*Quarter*	Quarterly mean prices $	Mean index
98	I	528.9	103.0
98	II	474.5	92.4
98	III	524.1	102.0
98	IV	527.0	102.6
99	I	474.4	92.4
99	II	546.0	106.3
99	III	562.4	109.5
99	IV	590.1	114.9
00	I	580.5	113.0
00	II	569.4	110.9
00	III	587.0	114.3
00	IV	561.4	109.3
01	I	535.7	104.3
01	II	524.2	102.1
01	III	463.6	87.7
01	IV	471.6	89.2
02	I	447.3	84.6
02	II	436.0	82.4
02	III	353.4	66.8
02	IV	315.8	61.5

Table A1.2 Data set variabels

Variable	Description	type of measure
lnp	log of price	dollars
lnpix	log of sharpness	megapixels
lnsto	log of memory included	megabytes
movie	movie feature	0 - 1 variable
remote	remote control	0 - 1 variable
flash_ex	external flash	0 - 1 variable
manfocus	manual focus	0 - 1 variable
zoomo	optical zoom	scale of optical magnification
zoomd	digital zoom	scale of digital magnification
USB	usb connection	0 - 1 variable
serial	serial connection	0 - 1 variable
bat_re	battery recharger	0 - 1 variable
type	type of camera	compact, ultacomp, SLR-type
multires	choices of various resolutions	number, or '0 - 1 variable
ISO	number of different iso	number, or '0 - 1 variable
manufac	manufacturer	18 manufacturers

TABLE A1.3. CLASSIFICATION METHOD

Year	Quarter	Only Manufacturer (1)	Manufacturer and Pixel-group (2)	2 + Manual focus
1998	I	103.4	105.4	106.0
1998	II	92.3	93.3	91.9
1998	III	96.4	95.4	100.4
1998	IV	108.0	105.8	101.7
1999	I	94.6	92.0	88.9
1999	II	116.5	108.6	97.9
1999	III	114.8	106.0	95.3
1999	IV	122.2	104.9	92.6
2000	I	112.8	96.0	85.4
2000	II	109.8	91.3	82.8
2000	III	111.3	85.3	74.9
2000	IV	110.9	77.3	67.9
2001	I	102.0	66.5	59.0
2001	II	102.5	63.9	56.4
2001	III	89.6	55.8	49.9
2001	IV	91.9	54.4	49.5
2002	I	83.9	47.7	43.1
2002	II	80.1	44.1	40.4
2002	III	63.3	36.0	33.5
2002	IV	58.7	33.9	31.5

Appendix 3. Some regression results

Time indicator Model 4 estimation results. The model is estimated as

$$\ln\left(\hat{p}_i^t\right) = \hat{\beta}^t x_i^t =$$

$$\hat{\alpha} + \hat{\beta}_1 LNPIX + \hat{\beta}_2 LNSTO + \hat{\beta}_3 MANFOCUS + \hat{\beta}_4 ZOOMO + \hat{\beta}_5 FLASH\_EX + \sum_{t=10}^{28} \hat{\delta}^t Q_t$$

The SAS System 09:33 Monday, May 12, 2001

The REG Procedure
Model: Model 4
Dependent Variable: lnp

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	24	170.35142	7.09798	160.64	<.0001
Error	938	41.44658	0.04419		
Corrected Total	962	211.79800			

Root MSE	0.21020	R-Square	0.8043
Dependent Mean	6.05545	Adj R-Sq	0.7993
Coeff Var	3.47133		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	6.56490	0.07584	86.57	<.0001
lnpix	ln(PIXEL)	1	0.50031	0.02129	23.50	<.0001
Manfocus	Manfocus	1	0.13551	0.01722	7.87	<.0001
ZoomO	Opt Zoom	1	0.06494	0.00562	11.56	<.0001
lnsto	ln(STORAGE)	1	0.12207	0.01539	7.93	<.0001
Flash_ex	Ext flash	1	0.11777	0.01872	6.29	<.0001
Q10		1	-0.16623	0.09449	-1.76	0.0789
Q11		1	-0.34812	0.08709	-4.00	<.0001
Q12		1	-0.49367	0.08241	-5.99	<.0001
Q13		1	-0.69502	0.08191	-8.48	<.0001
Q14		1	-0.77254	0.08240	-9.38	<.0001
Q15		1	-0.80633	0.07965	-10.12	<.0001
Q16		1	-0.90879	0.08450	-10.76	<.0001
Q17		1	-0.93108	0.07829	-11.89	<.0001
Q18		1	-0.97396	0.08171	-11.92	<.0001
Q19		1	-1.07752	0.08038	-13.41	<.0001
Q20		1	-1.15870	0.08123	-14.26	<.0001
Q21		1	-1.29944	0.08044	-16.15	<.0001
Q22		1	-1.37906	0.08708	-15.84	<.0001
Q23		1	-1.50989	0.08095	-18.65	<.0001
Q24		1	-1.52072	0.08508	-17.87	<.0001
Q25		1	-1.59569	0.08398	-19.00	<.0001
Q26		1	-1.65902	0.08160	-20.33	<.0001
Q27		1	-1.83195	0.07983	-22.95	<.0001
Q28		1	-1.90307	0.08005	-23.77	<.0001

Appendix 4. *Model 3 and 4 coefficients*

Models 3 and 4 estimated from total and pairwise pooled data

Table A4.1 Parameter estimates for pairwise dummy model and full dummy model 3

Year	Qrt	Intercept	lnpix	Manfocus	ZoomO	Time D
98	I					
98	II	6.81	0.65	0.07	0.06	-0.167
98	III	6.52	0.52	0.13	0.07	-0.135
98	IV	6.33	0.42	-0.06	0.08	-0.079
99	I	6.25	0.43	0.04	0.07	-0.153
99	II	6.09	0.49	0.05	0.08	-0.025
99	III	6.05	0.54	0.14	0.07	-0.047
99	IV	5.99	0.53	0.21	0.07	-0.078
00	I	5.89	0.58	0.16	0.08	-0.037
00	II	5.84	0.60	0.15	0.08	-0.045
00	III	5.83	0.54	0.17	0.07	-0.085
00	IV	5.76	0.55	0.14	0.07	-0.089
01	I	5.66	0.60	0.14	0.07	-0.141
01	II	5.51	0.61	0.16	0.06	-0.089
01	III	5.41	0.59	0.08	0.09	-0.110
01	IV	5.28	0.57	0.10	0.10	-0.015
02	I	5.30	0.51	0.21	0.08	-0.083
02	II	5.15	0.53	0.26	0.09	-0.054
02	III	5.13	0.53	0.23	0.08	-0.186
02	IV	5.03	0.46	0.23	0.08	-0.063
Full model		**6.69**	**0.53**	**0.18**	**0.08**	

Table A4.2 Parameter estimates for pairwise dummy model and full dummy model 4

Year	Qrt	Intercept	lnpix	lnsto	Manfocus	ZoomO	Flash_ex	Time D
98	I							
98	II	6.68	0.65	0.18	0.00	0.04	0.00	-0.153
98	III	6.46	0.51	0.06	0.10	0.07	-0.12	-0.153
98	IV	6.28	0.41	0.04	-0.06	0.08	-0.02	-0.088
99	I	6.17	0.42	0.06	-0.02	0.07	0.03	-0.184
99	II	5.97	0.47	0.08	0.07	0.05	0.12	-0.039
99	III	5.78	0.34	0.11	0.16	0.07	0.22	-0.019
99	IV	5.75	0.35	0.08	0.18	0.10	0.20	-0.095
00	I	5.55	0.47	0.13	0.12	0.11	0.13	-0.015
00	II	5.60	0.53	0.08	0.09	0.12	0.09	-0.054
00	III	5.52	0.46	0.11	0.08	0.13	0.06	-0.089
00	IV	5.37	0.44	0.16	0.06	0.11	0.10	-0.074
01	I	5.40	0.55	0.14	0.11	0.04	0.11	-0.143
01	II	5.26	0.57	0.14	0.14	0.02	0.13	-0.087
01	III	5.22	0.64	0.09	0.05	0.06	0.11	-0.134
01	IV	5.11	0.68	0.07	0.04	0.06	0.13	-0.027
02	I	5.15	0.78	0.01	0.14	0.04	0.09	-0.094
02	II	4.88	0.68	0.10	0.15	0.06	0.08	-0.071
02	III	4.79	0.56	0.15	0.13	0.06	0.12	-0.170
02	IV	4.80	0.49	0.11	0.18	0.05	0.10	-0.073
Full model		**6.56**	**0.50**	**0.12**	**0.14**	**0.12**	**0.06**	

Table A4.3 Parameter estimates for full hedonic model 3

Year	Qrt	Intercept	lnpix	Manfocus	ZoomO
98	I	6.75	0.58		0.07
98	II	6.72	0.71	0.06	0.05
98	III	6.35	0.47	0.13	0.07
98	IV	6.25	0.38	-0.11	0.08
99	I	6.09	0.48	0.13	0.07
99	II	6.08	0.51	0.00	0.08
99	III	5.99	0.54	0.22	0.07
99	IV	5.91	0.51	0.20	0.07
00	I	5.84	0.61	0.14	0.08
00	II	5.79	0.57	0.17	0.08
00	III	5.77	0.53	0.16	0.07
00	IV	5.66	0.59	0.11	0.08
01	I	5.51	0.60	0.17	0.06
01	II	5.42	0.65	0.16	0.06
01	III	5.28	0.60	0.06	0.10
01	IV	5.30	0.52	0.18	0.09
02	I	5.22	0.50	0.24	0.08
02	II	5.06	0.55	0.28	0.09
02	III	4.98	0.53	0.20	0.08
02	IV	5.01	0.39	0.26	0.07

Table A4.4 Parameter estimates for full hedonic model 4

Year	Qrt	Intercept	lnpix	Manfocus	ZoomO	Insto	Flash_ex
98	I	6.63	0.59		0.05	0.16	
98	II	6.58	0.71	-0.01	0.03	0.19	
98	III	6.36	0.48	0.12	0.08	0.00	-0.05
98	IV	6.12	0.36	-0.11	0.09	0.08	-0.02
99	I	6.01	0.48	0.04	0.07	0.05	0.04
99	II	5.80	0.47	0.12	0.03	0.16	0.14
99	III	5.79	0.33	0.18	0.10	0.06	0.23
99	IV	5.58	0.34	0.18	0.10	0.12	0.18
00	I	5.50	0.52	0.09	0.12	0.13	0.11
00	II	5.72	0.58	0.09	0.13	-0.02	0.05
00	III	5.28	0.38	0.06	0.13	0.19	0.08
00	IV	5.40	0.53	0.07	0.07	0.14	0.10
01	I	5.26	0.57	0.15	0.02	0.14	0.13
01	II	5.19	0.57	0.13	0.03	0.14	0.13
01	III	5.08	0.66	0.01	0.07	0.08	0.10
01	IV	5.15	0.73	0.07	0.05	0.02	0.18
02	I	5.04	0.80	0.17	0.04	0.00	0.00
02	II	4.72	0.62	0.14	0.07	0.15	0.12
02	III	4.67	0.54	0.13	0.05	0.15	0.11
02	IV	4.88	0.43	0.24	0.05	0.06	0.07

Table A4.5 Average quality characteristics

Year	Qrt	lnp	lnpix	Manfocus	ZoomO	lnsto	Flash_ex
98	I	6.22	-0.96		0.59	0.85	
98	II	6.09	-0.93	0.09	0.75	0.82	
98	III	6.20	-0.49	0.06	0.96	1.22	0.06
98	IV	6.23	-0.23	0.15	1.08	1.49	0.04
99	I	6.11	-0.14	0.14	1.30	1.56	0.14
99	II	6.27	0.10	0.28	1.95	1.84	0.34
99	III	6.27	0.17	0.26	1.95	1.89	0.33
99	IV	6.32	0.33	0.35	2.28	2.06	0.42
00	I	6.29	0.35	0.37	2.18	2.02	0.34
00	II	6.28	0.39	0.38	2.36	2.03	0.35
00	III	6.30	0.53	0.51	2.43	2.21	0.41
00	IV	6.26	0.63	0.53	2.39	2.23	0.33
01	I	6.20	0.73	0.61	2.49	2.33	0.33
01	II	6.21	0.84	0.69	2.58	2.39	0.46
01	III	6.07	0.83	0.65	2.75	2.38	0.34
01	IV	6.10	0.89	0.60	3.03	2.36	0.34
02	I	6.01	0.92	0.60	2.80	2.34	0.26
02	II	5.98	0.97	0.60	2.66	2.49	0.22
02	III	5.76	0.90	0.55	2.65	2.50	0.18
02	IV	5.66	0.87	0.56	2.59	2.46	0.20

Table A4.6: Quality correction factors

Year	Qrt	Model 2P	Model 2Full	Model 4P	Model 4Full	Model 2	Model 4
98	I						
98	II	-0.04	-0.05	-0.02	-0.05	-0.03	-0.02
98	III	-0.24	-0.25	-0.26	-0.30	-0.27	-0.31
98	IV	-0.11	-0.16	-0.12	-0.18	-0.12	-0.13
99	I	-0.06	-0.06	-0.06	-0.09	-0.06	-0.06
99	II	-0.18	-0.20	-0.20	-0.26	-0.18	-0.20
99	III	-0.03	-0.03	-0.02	-0.04	-0.03	-0.03
99	IV	-0.13	-0.13	-0.14	-0.16	-0.13	-0.14
00	I	-0.01	-0.01	0.02	0.01	-0.01	0.02
00	II	-0.04	-0.04	-0.05	-0.05	-0.04	-0.05
00	III	-0.10	-0.10	-0.10	-0.12	-0.10	-0.10
00	IV	-0.06	-0.06	-0.04	-0.05	-0.06	-0.04
01	I	-0.08	-0.07	-0.08	-0.08	-0.08	-0.08
01	II	-0.08	-0.08	-0.10	-0.09	-0.09	-0.10
01	III	-0.01	0.00	0.01	0.00	0.00	0.02
01	IV	-0.06	-0.05	-0.06	-0.06	-0.06	-0.06
02	I	0.00	0.00	-0.01	0.02	0.00	-0.01
02	II	-0.01	-0.01	-0.04	-0.02	-0.01	-0.03
02	III	0.05	0.05	0.05	0.04	0.05	0.05
02	IV	0.02	0.02	0.02	0.03	0.02	0.02

These are derived as $\sum \hat{\beta}_1 '\left(\overline{x}_1^{t-1} - \overline{x}_1^t\right) + \ldots + \hat{\beta}_K '\left(\overline{x}_K^{t-1} - \overline{x}_K^t\right)$

Table A4.7 Quality correction factors

Year	Qrt	Model 2P	Model 2Full	Model 4P	Model 4Full	Model 2	Model 4
98	I	117.7	118.5	117.6	121.5	119.1	120.2
98	II	113.4	115.0	114.9	117.6	115.2	117.5
98	III	89.1	89.9	88.7	87.8	87.7	86.3
98	IV	79.8	76.7	78.8	73.1	78.0	76.0
99	I	75.5	71.8	74.0	67.0	73.7	71.4
99	II	63.3	58.9	60.6	51.6	61.7	58.4
99	III	61.3	57.0	59.4	49.8	59.7	56.9
99	IV	53.7	50.0	51.5	42.3	52.4	49.4
00	I	53.4	49.7	52.4	42.7	52.1	50.3
00	II	51.2	47.9	49.9	40.8	49.9	47.9
00	III	46.4	43.4	45.0	36.3	45.2	43.2
00	IV	43.8	41.0	43.4	34.6	42.6	41.5
01	I	40.6	38.1	39.9	31.8	39.5	38.2
01	II	37.3	35.2	36.1	29.0	36.2	34.6
01	III	37.1	35.2	36.6	29.0	36.2	35.2
01	IV	35.0	33.6	34.5	27.5	34.2	33.2
02	I	35.0	33.6	34.2	28.0	34.2	32.9
02	II	34.6	33.1	33.0	27.3	33.8	31.8
02	III	36.4	34.7	34.6	28.6	35.6	33.5
02	IV	37.0	35.4	35.3	29.3	36.2	34.1

In August 2004 Statistics Finland hosted the 8th meeting of the Ottawa Group, the international working group on price indices. This publication contains the post-conference final versions of the papers that were presented in the meeting as well as the session summaries prepared by the session moderators. In addition to this publication, these papers will in the future also be available through the Ottawa Group web-site www.ottawagroup.org.

Tilastokeskus, myyntipalvelu
PL 4C
00022 TILASTOKESKUS
puh. (09) 1734 2011
faksi (09) 1734 2500
myynti@tilastokeskus.fi
www.tilastokeskus.fi

Statistikcentralen, försäljning
PB 4C
00022 STATISTIKCENTRALEN
tfn (09) 1734 2011
fax (09) 1734 2500
myynti@stat.fi
www.stat.fi

Statistics Finland, Sales Services
P.O.Box 4C
FI-00022 STATISTICS FINLAND
Tel. +358 9 1734 2011
Fax +358 9 1734 2500
myynti@stat.fi
www.stat.fi

ISSN 1239–3800
= Reviews
ISBN 952–467–530–7
Product number 91185