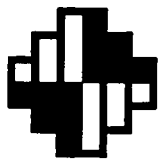


P15,3



## Tilastokeskus

Tekijä

Päiväys

nro

Christian Starck

3.11.1988

117

Väestölaskenta 1985

KÄSITTELYVIRHESELVITYS

Tilastokirjasto

TILASTOKIRJASTO



136 003 6440

## VUODEN 1985 VÄESTÖLASKENTA: KÄSITTELYVIRHESELVITYS

1. Johdanto .....	1
2. Lomakeaineiston käsittely .....	2
3. Käsittelyvirheselvityksen tutkimusasetelma .....	7
- tallennusvirheselvitys	
- tallennetun aineiston käsittelyvirheselvitys	
4. Tallennusselvityksen tulokset .....	10
5. Muuttujakohtainen vertailu .....	14
- vertailun periaatteet ja tunnuslukujen laskeminen	
- pääasiallinen toiminta	
- työssäolo- ja työttömyyskuukaudet	
- ammatissatoimivuus	
- ammattiasema	
- ammatti	
- elinkeino (toimiala)	
- työpaikan sijaintikunta	
- ns. liikkuvuustyyppi	
- työpaikan juridinen muoto	
6. Vertailu vuoden 1980 laskennan käsittely- virheselvityksen tuloksiin .....	32
7. Käsittelyvirheiden seuraukset .....	35
8. Lopuksi .....	37
Kirjallisuusluettelo .....	40
Liitteet	
- liitetaulukot	
- laskentalomake (esitäytetty)	
- tallennusohjeet	

## 1. JOHDANTO

Tässä raportissa kerrotaan niistä tuloksista joita saatiin, kun parintuhannen kertaalleen käsitellyn väestölaskentalomakkeen tiedot tallennettiin uudelleen ja näitä tietoja verrattiin alkuperäisiin tallennuskenttiin. Raportissa kerrotaan myös selvituksesta, jossa uuden tallennustiedoston pohjalta tehtyjä muuttujia verrattiin varsinaisen laskenta-aineiston vastaaviin tietoihin.

Kyseessä olevan käsittelyvirheselvityksen keskeisimpänä tavoitteena on valottaa tallennuksen ja ns. käsittelysysteemien luotettavuutta. Vuoden 1985 väestölaskennan yhteydessä otettiin käyttöön lukuisia uusia menetelmiä: tallennus ilman esitarkistusta, ammatti- ja työpaikkatietojen esitäyttö, ns. automaattinen kooditus sekä vuorovaikutteinen päätekooditus. Olivatko uudet ratkaisut käytännössä toimivia? Mikäli käsittelyselvitystä varten poimittu otos toisella kierroksella tuottaisi lukuisia poikkeavia muuttuja-arvoja, tämä tietäisi huonoa arvosanaa laskenta-aineiston käsittelyratkaisuille.

Raportissa ei kuitenkaan tarkemmin selvitetä, mitkä osat käsittelysysteemistä mahdollisesti ovat omiaan tuottamaan virrehajontaa. Käsittelysysteemi käsitteään lähinnä "mustana laatikkona" jonka sisäosiin ei kajota. Lomaketietojen käsittelyä arvioidaan kokonaisuutena.

Käsittelyvirheselvityksen toisena tavoitteena on tuottaa tietoa väestölaskentatilastojen laadusta. Mitä enemmän samoista lomaketiedoista samoilla määritelmillä tehdyt tilastot poikkeavat toisistaan, sitä varauksellisemmin pitää luonnollisesti suhtautua laskenta-aineiston pohjalta laadittuihin taulukoihin.

Se, että uusintatutkimus tuottaa lähes samanlaisia jakaumia on välttämätön, mutta ei riittävä ehto tietojen "oikeellisuudelle". Onhan mahdollista, että nyt käytössä olleet ratkaisut systemaattisesti tuottavat jollakin tavalla vääristyneitä tietoja. Tietojen paikkansapitävyyttä ei kuitenkaan tutkita tässä, vaan erillisessä vuoden 1985 väestölaskennan luotettavuustutkimuksessa.

## 2. LOMAKEAINEISTON KÄSITTELY

Vuoden 1985 väestölaskennassa lähetettiin lomake kaikille 15-74-vuotiaille, joilla oli 'lainmukainen kotipaikka Suomessa'. Laskentalomakkeella oli yhteensä kahdeksan kysymystä:

1. henkilötunnus
2. toiminta laskentaviikolla 11.11.-17.11.1985
3. työpaikan nimi
4. työpaikan toimiala
5. työpaikan sijaintikunta
6. työpaikan lähiosoite
7. ammatti
8. työssäolo- ja työttömyyskuukaudet

Lomakkeille esitätettiin vastaanottajan nimi ja osoite, henkilötunnus sekä väestön keskusrekisterin mukainen ammattinimike vuodelta 1985 (mikäli se automaattisen koodituksen kannalta oli riittävän yksiselitteinen). Lisäksi esitätettiin vuoden 1980 väestölaskennan työpaikkatiedot: työpaikan nimi ja osoite (tietyin poikkeuksin).

Palautetut väestölaskentalomakkeet toimitettiin tallennuksesta vastaavalle yritykselle sellaiseenaan, ilman edeltävää järjestämis-, tarkistus- ja kooditustyötä. Lomaketietojen tallennus suoritettiin atk-kirjoituksena siten, että tallennetun tietueen voidaan katsoa olevan 1/1-kopio väestölaskentailmoituksesta. Tallennus koski yhteensä 17 tietoa, joista kolme oli tietueen tunnistustietoja (eränumero, nipun numero, henkilötunnus). Muut tiedot olivat seuraavat:

Kys.	Tieto	Pituus	Tunniste
2	Toiminta laskentaviikolla	10	TOI
3	Työpaikan nimi	36	TYNI
4	Toimiala	36	TOLA
4	Toimiala maa-, metsä- tai karjatalous	1	MAA
4	Metsätyökuukaudet	2	METS
5	Työpaikan kuntarasti	3	LKRAS
5	Työpaikan kunta	18	TPKUN
6	Lomakkeen osoiterasti	3	LORAS
6	Työpaikan lähiosoite	36	TPOS
7	Ammatti	36	LAMMA
8	Kokopäivätyökuukaudet	2	KTYÖ
8	Osa-aikatyökuukaudet (väh. 20 h 2)	OSTYÖ	
8	Osa-aikatyökuukaudet (alle 20 h/v)	2	OSA
8	Työttömyyden kesto	2	TYKE

Tallennuskenttien pituus oli yhteensä 205 merkkiä. Keskimäärin tallennettiin 45.3 merkkiä/lomake.

Osa tallennustiedoista oli aakkostietoa, osa taas numeerista tietoa. Tiettyihin tallennuskenttiin kone hyväksyi vain ennalta määrätyn arvojoukon. Tallennusohjeet löytyvät liitteestä 2.

Tallennuskenttien tunnisteita (TOI, TYNI jne.) käytetään jatkossa silloin, kun nimenomaan viitataan tallennettuun merkkijonoon, eikä niihin muuttujiin, joita muodostettiin tallennettuja tietoja hyväksikäyttämällä.

Lomaketietojen jatkokäsittely jaettiin viiteen osasysteemiin, joiden keskeinen sisältö oli lyhyesti seuraava:

#### I Lomakeaineiston peruskäsittely

Tallennetut kooditiedot tarkistettiin, suoritettiin mm. pääasiallisen toiminnan päättely, esikäsiteltiin työpaikka- ja ammattitietoja ja määriteltiin täydennystiedustelutarve.

Työpaikan sijaintikunnan kooditusta varten oli muodostettu konekielinen tiedosto, johon kuntien "oikeiden" nimien ja kuntakoodien lisäksi vietiin joitakin muita paikannimiä, joita ennakolta oletettiin tiedonantajan käyttävän kunnan nimen asemesta.

Peruskäsittely suoritettiin mahdollisimman pitkälle koneellisesti. Päätekäsittelyssä esikäsiteltiin noin 10 % tietueista.

#### II Työpaikkatietojen käsittely

Osasysteemin tavoitteena oli tuottaa lomakkeella ilmoitetuista selväkielisistä työpaikkatiedoista koodimuodossa työllisen väestön

- työpaikan toimiala
- työpaikan sijainti (kunta, taajama, kunnan osa-alue ja karttakoordinaatit)
- työpaikan juridinen muoto

Tietojen käsittelyn kannalta työpaikkatiedot jakautuivat kolmeen ryhmään:

1. Työpaikka oli pysynyt samana kuin lomakkeelle esitetytty (vuodelta 1980).

Vuoden 1980 laskennan toimipaikkatiedosto sisälsi nimi- ja osoitetietojen lisäksi tiedot työpaikan toimialasta ja juridisesta muodosta koodimuodossa sekä sijainnin karttakoordinaa-

tein, joten kaikki kyseisten henkilöiden työpaikkatiedot saatiin suoraan toimipaikkatiedostosta eräajona ilman manuaalikäsittelyä.

2. Lomakkeella ilmoitettu työpaikka sijaitsi kiinteässä toimipaikassa.

Tietojen koneellinen ja manuaalinen jatkokäsittely perustui sijaintikuntatietoa lukuunottamatta toimipaikkatiedoston hyväksikäyttöön. Toimipaikkatietueille muodostettiin nimestä (5 merkkiä) ja osoitteesta (6-8 merkkiä) tietyillä säännöillä tunnistusta varten vertailukenttä.

Tallennetuista työpaikan nimi- ja osoitetiedoista muodostettiin vastaava vertailukenttä. Mikäli työpaikka yksikäsitteisesti tunnistui koneellisesti toimipaikkatiedoston toimipaikkaan samassa kunnassa, siirrettiin henkilölle tämän toimipaikan toimipaikkatunnus.

Mikäli työpaikka ei tunnistunut koneellisessa vertailussa mihinkään toimipaikkatiedoston toimipaikkaan, tehtiin henkilön toimipaikkatiedoista koneellisesti tiedostoon toimipaikkali säys, jonka toimipaikkatunnus liitettiin henkilön tietueelle. Kaikki näin lisätyt toimipaikat tarkistettiin päätetyönä.

Mikäli työpaikka tunnistui useampaan kuin yhteen toimipaikkaan, työpaikkatiedot siirtyivät manuaaliseen pätekkäsittelyyn.

Koska työpaikkatiedon käsittely perustui sen mahdollisimman hyvään tunnistuvuuteen toimipaikkatiedostosta, tiedosto pyrittiin saamaan mahdollisimman kattavaksi. Siksi vuoden 1980 väestölaskennan toimipaikkatiedostoa täydennettiin mm. Tilastokeskuksen yritystietorekisterin toimipaikoilla, väestön keskusrekisterin toimintatiedoilla sekä Kunnallisen Eläkelaitoksen toimintayksikkörekisteristä saaduilla tiedoilla.

Lomakeaineiston käsittelyn päätyttyä toimipaikkatiedosto sisälsi 0,96 milj. tietuetta. Koneellisen ja manuaalisen tuplatietueiden poiston jälkeen toimipaikkatietueita oli kaikkiaan 0,19 milj.

Kun toimipaikkatiedoston käsittely oli saatu päätökseen, siirrettiin henkilöille, joiden työpaikka sijaitsi toimipaikkatiedoston toimipaikassa, henkilötietueelle viedyn toimipaikka-

tunnuksen avulla työpaikan toimiala, juridinen muoto ja sijaintitiedot.

Lomaketietojen käsittelyn alussa työpaikkatietojen koneellinen koodiutuvuus oli noin 35 %. Viimeisessä käsittelyerässä se oli jo 65 %. Automaattisesti koodiutui kaikkiaan 0,7 milj. henkilön työpaikkatiedot.

3. Lomakkeella ilmoitettu työpaikka ei sijainnut kiinteässä toimipaikassa.

Tähän ryhmään kuuluivat kaikki ne, joiden työ oli luonteeltaan liikkuvaa sekä myös esim. kaikki kotona työskentelevät. Liikkuvaa työtä tekeviltä puuttui aina osoitetieto ja ammatinharjoittajalta työpaikan nimi. Näihin ryhmiin kuuluneiden henkilöiden toimiala ja juridinen muoto kooditettiin kokonaan manuaalisena päätetyönä.

### III Ammattitiedon käsittely

Osasysteemin tavoitteena oli tuottaa lomakkeella ilmoitetusta selväkielisestä ammattitiedosta työllisen väestön ammattitiedot koodimuodossa.

Ammattitiedon käsittelyssä sovellettiin mahdollisimman pitkälle automatisoituja menetelmiä. Automaattinen kooditusjärjestelmä koostui ammattinimikkeistä, ammattikoodeista ja eräistä aputiedoista muodostetusta ammattisanakirjasta sekä sen käsittelyyn liittyvästä ohjelmistosta. Ammattisanakirjan perustan muodostivat kaikki Tilastokeskuksen ammattiluokituksen suomen- ja ruotsinkieliset nimikkeet. Kaiken kaikkiaan ammattisanakirjassa oli noin 16 000 nimikettä. Samasta nimikkeestä saattoi esiintyä useita eri kirjoitusasuja. Selväkielisenä lomakkeelta tallennetut ammattitiedot siirtyivät ammatin automaattiseen eräajokooditukseen, jossa koodiutui kaikkiaan noin 1,2 milj. ammattinimikettä. Niiden noin 0,48 milj. ammattinimikkeen manuaalikäsitteily, jotka eivät koodiutuneet automaattisesti eräajossa, suoritettiin päätetyönä.

Ammattisanakirjaan ei tehty systemaattisesti lisäyksiä käsittelyn aikana.

#### IV Lomaketietojen täydentäminen rekisteri- tiedoilla

Tavoitteena oli tietokohtaisen peittävyden nostaminen mahdollisimman korkeaksi.

Täydennyksessä käytetyt rekisterit olivat väestön keskusrekisteri, Eläketurvakeskuksen rekisteri, Kansaneläkelaitoksen rekisteri, yritysrekisteri, verotuksen rekisterit, kuntien toimintayksikkörekisteri sekä julkisen sektorin palkkarekisterit. Rekisteritäydennyksen avulla saatiin aineistoa täydennettyä 56 000 työllisellä, joista saatiin seuraavat tiedot:

- ammattiasema
- juridinen muoto
- toimiala
- työpaikan sijaintikunta
- ammatissatoimivuus

#### V Lomaketietojen lopputarkistus

Tarkistettiin ja korjattiin edeltävissä osasysteemeissä osittain toisistaan riippumattomina muodostettujen kooditietojen loogiset ristiriidat.

Edellä luetellut osasysteemit jakautuivat kukin erikseen täysin koneellisesti toteutettuun eräajosysteemiin ja päätetyönä toteutettuun manuaalikäsitteleyksysteemiin.

Tarkempi kuvaus käsittelyn osasysteemeistä löytyy vuoden 1985 väestölaskennan systeemis suunnitelmasta (Tilastokeskus, VL 85, 3.5.1985), loppuraportista (Tilastokeskus, Raportti 1988:7) sekä lomakekeruun loppuraportista (Tilastokeskus, Muistio 116, 1988).

### 3. KÄSITTELYVIRHESELVITYKSEN TUTKIMUSASETELMA

#### *Tallennusselvitys*

Tallennusselvitystä varten tallennettiin uudestaan niiden henkilöiden lomaketiedot, jotka olivat olleet mukana Tilastokeskuksen työvoimatutkimuksen marraskuun 1985 otoksessa (ensimmäinen rotaatio-ryhmä). Tämä otos (N = 2 344, joista vastanneita 2 194) katsottiin sopivaksi monistakin syistä: se oli satunnaisotos, se oli sopivan kokoinen ja - ennen kaikkea - kyseistä perusjoukkoa oli tarkoitus tutkia laskennan muissakin luotettavuusselvityksissä. Valitsemalla tämä otos käsittelyvirheselvityksen perusjoukoksi toivottiin saatavan monipuolinen ja tarkka kuva laskenta-aineistossa esiintyvistä virheistä.

Lähes kaikkien otoksessa olevien henkilöiden laskentalomakkeet löytyivät. Muutamasta henkilöstä löytyi jopa kaksi lomaketta: varsinaisen tiedustelun lomake ja täydennystiedustelulomake. Tallennusselvityksessä oli mukana yhteensä 2 210 lomaketta.

Tallennusselvitys tehtiin seuraavalla tavalla:

1. Lomakkeilla olevat tiedot tallennettiin uudestaan. Tallennusohjeet olivat samat kuin varsinaisen tallennuksen yhteydessä.
2. Uutta tallennustiedostoa vertailtiin koneellisesti kenttä kentältä alkuperäisiin tallennustietoihin.
3. Poikkeavista tiedoista tulostettiin luettelo, josta saatiin selville ne lomakkeet ja kentät, joissa esiintyi poikkeavia tietoja. Poikkeavien tallennuskenttien tiedot tulostettiin kokonaisuudessaan. Tieto katsottiin poikkeavaksi, mikäli yksikin merkki tiedolle varatussa kentässä poikkesi alkuperäisestä tallennuksesta.
4. Kaikkia poikkeavia tietoja vertailtiin alkuperäisiin lomakkeisiin. Uusi tallennustieto korjattiin, mikäli se selvästi oli virheellinen. Mikäli tieto molemmilla tallennuskerroilla oli täysin sama, katsottiin, että se vastasi lomakkeella annettua tietoa.
5. Uutta - ja nyt mahdollisimman virheetöntä - tallennustiedostoa verrattiin taas koneellisesti

alkuperäiseen tiedostoon, jonka jälkeen uusi luettelo poikkeavuuksista tulostettiin. Samalla laskettiin myös poikkeavien kenttien lukumäärät.

Varsinainen tallennusselvitys tehtiin siis siten, että mahdollisimman virheettömiä tallennustietoja verrattiin alkuperäisiin tallennustietoihin.

### *Tallennetun aineiston käsittelyvirheselvitys*

Uudelleen tallennettu aineisto muokattiin käyttäen muuttujia - mahdollisimman samoja käsittelysääntöjä kuin varsinaisen laskennan yhteydessä. Aineistolle tehtiin ns. peruskäsittely, jonka jälkeen ne tietueet, joilla oli selväkielisiä työpaikka- ja ammattitietoja siirtyivät ns. työpaikka- ja ammattisysteemeihin, joissa työllisille pyrittiin antamaan oikeat ammatti-, toimiala- ym. koodit. Kaikille tietueille tehtiin myös joitakin lopputarkistuksia. Kooditus tapahtui sekä koneellisesti että manuaalisesti.

Uudelleen tallennetun aineiston käsittely ei kuitenkaan kaikkien tietojen osalta täysin vastannut aikaisempaa käsittelyä. Puuttuvia lomaketietoja ei täydennetty tiedoilla hallinnollisista rekistereistä. Täydennykset olisivat koskeneet vain muutamaa tietuetta ja katsottiin, että suurta työmäärää vaativasta toimenpiteestä ei olisi ollut vastaavaa hyötyä. Käytännön syistä täydennystiedustelulomaketta ei voitu lähettää niille muutamille henkilöille, joiden lomake tallennettujen tietojen perusteella luokiteltiin puutteellisesti täytetyksi. Myös joitakin koneellisia lopputarkistuksia jäi uusintatutkimuksen yhteydessä tekemättä.

Aineistokäsittelyn jälkeen laskettiin muuttujakohtaisia tunnuslukuja samojen ja poikkeavien koodiarvojen osuuksista. Alkuperäisten ja uusien muuttujien suoria jakaumia verrattiin myös keskenään. Vertailussa oli mukana yhteensä kolmetoista muuttujaa. Havaintoja oli yhteensä 2 188 eli muutama vähemmän kuin tallennusselvityksessä. Tämä johtui siitä, että sama henkilö saattoi olla mukana tallennusselvityksessä kahdella lomakkeella.

Se, että joitakin lopputarkistuksia uusintatutkimuksen yhteydessä jäi tekemättä on jonkin verran vaikeuttanut havaittujen erojen tulkintaa: on ollut vaikeata tietää johtuvatko poikkeavuudet tallennuksesta, käsittelyvirheistä, vaiko siitä, että jokin lopputarkistus syystä tai toisesta jäi tekemättä.

Uusintatutkimuksen tuloksia ei näin voida pitää määritelmällisesti 'oikeina'. Tämä ei toisaalta johdu pelkästään siitä, ettei kaikkia lopputarkistuksia tehty tai siitä, että huolimattomuusvirheitä varmaankin esiintyi myös uusintatutkimuksen yhteydessä. Tärkeä syy on myös se, että erityisesti ammatti- ja toimialatietojen kohdalla monta kertaa on vaikeata löytää yksiselitteisesti 'oikea' koodi. Näihin ongelmiin palataan vielä myöhemmin.

## 4. TALLENNUSSELVITYKSEN TULOKSET

Poikkeavuuksia tiedostojen välillä esiintyi paljon: taulukosta 1 voidaan nähdä, että itse asiassa vain 62 % lomakkeista tallennettiin molemmilla kerroilla täysin samalla tavalla. Tämä johtui kuitenkin lähinnä siitä, että aakkostietojen kohdalla käytettiin joskus lyhenteitä (mikä oli sallittua), joskus ei (mikä myös oli sallittua). Mikäli tarkastellaan vain kenttiä, joissa oli numeerisia tietoja (lomakkeella numeroita tai rasteja), identtisten tietueiden osuus oli 95 %.

Taulukko 1: Lomakekohtaisia tietoja poikkeavuuksista

Poikkeamaa	Lomakkeita yhteensä			Aakkos- tiedot %	Num.- tiedot %	Kaikki tiedot %
	Aakkos- tiedot	Num.- tiedot	Kaikki tiedot			
0	1 419	2 097	1 380	64.2	94.9	62.4
1	458	77	451	20.7	3.5	20.4
2	238	18	253	10.8	0.8	11.4
3	75	6	84	3.4	0.3	3.8
4	19	5	18	0.9	0.2	0.8
5	1	5	9	0.0	0.2	0.4
6+	-	2	15	-	0.1	0.7
Yhteensä	2 210	2 210	2 210	100.0	100.0	100.0

Mikäli osuudet lasketaan niistä kentistä, joihin uusintatallennuksen yhteydessä ylipäänsä oli tallennettu jotakin (9 687 kenttää), identtisten kenttien osuus kaikista kentistä oli 85 %. Vastaava osuus numeeristen kenttien kohdalla oli 97 % ja aakkostietojen kohdalla 33 %.

**Taulukko 2: Poikkeavien tallennustietojen määrä ja osuus**

Kysymys	Tieto	Tall. tiet.(a)	Yht. (b)	Poikkeavia tiet.(c)	(c/a) %	(c/b) %
2	TOI (n)	2153	2210	22	1.0	0.9
3	TYNI (a)	878	2210	289	33.0	13.1
4	MAA (n)	119	2210	11	9.2	0.5
4	METS (n)	68	2210	15	22.0	0.7
4	TOLA (a)	714	2210	186	26.0	8.4
5	LKRAS (n)	1050	2210	33	3.1	1.5
5	TPKUN (a)	287	2210	57	19.9	2.6
6	LORAS (n)	700	2210	59	8.4	2.7
6	TPOS (a)	850	2210	383	45.0	17.3
7	LAMMA (a)	1036	2210	325	31.4	14.7
8	KTYÖ (n)	1504	2210	25	1.7	1.1
8	OSTYÖ (n)	91	2210	10	11.0	0.5
8	OSA (n)	72	2210	6	8.3	0.3
8	TYKE (n)	165	2210	7	4.2	0.3
	Num. tiet.	5922	19890	188	3.2	0.6
	Aakk.tiet.	3765	11050	1240	32.9	19.4
	Yhteensä	9687	30940	1428	14.7	4.6

(n) = numeerinen tieto  
(a) = aakkostieto

Taulukossa 2 ensimmäinen prosenttisarake on laskettu siten, että peruslukuna on niiden tietueiden määrä, joihin uusintatallennuksen yhteydessä tallennettiin tietoa. Näin laskettuna osuudet kuvaavat parhaiten itse tallennustyön luotettavuutta. Toinen prosenttisarake antaa tietoa siitä, missä määrin uudelleen tallennettu tieto poikkesi alkuperäisestä aineistosta, kun myös tyhjät kentät otetaan huomioon. Peruslukuna on tällöin käytetty kaikkien tietueiden määrää (N = 2 210). Poikkeavuuksien suhteellinen osuus on näin laskettuna luonnollisesti pienempi.

Taulukosta 2 nähdään selvästi, kuinka aakkostietojen kohdalla poikkeavuudet olivat huomattavasti yleisempiä kuin numeeristen kenttien kohdalla. Silloin, kun numeeristen kenttien kohdalla esiintyi poikkeavuuksia, kysymyksessä oli usein lomake, joka tavalla tai toisella oli täytetty moniselitteisesti.

Numeerisen kentän METS (mestätyökuukaudet) kohdalla esiintyi verraten paljon poikkeavia tallennustietoja. Tämä johtunee ainakin osittain vastauskentän sijainnista lomakkeella - tallentajien mielestä vastausruutua olikin vaikeata havaita ja tieto saattoi jäädä tallentamatta.

Kuviossa 1 on esimerkkejä poikkeavuuksista.

NIPPU:	0867	
HTUNN:		
HUOM :	-	
TIETO:	TPKUN	HKI HELSINKI
-----		
NIPPU:	0871	
HTUNN:		
HUOM :	-	
TIETO:	LAMMA	PÄÄKASSANHOIT PÄÄKASSANHOITAJA
-----		
NIPPU:	0916	
HTUNN:		
HUOM :	-	
TIETO:	LAMMA	SÄHKÖASENT SÄHKÖASENTAJA
-----		
NIPPU:	0936	
HTUNN:		
HUOM :	-	
TIETO:	TYNI	JGNSUUN PGLIISILAITOS PGLIISILAITOS
-----		
NIPPU:	0967	
HTUNN:		
HUOM :	-	
TIETO:	TPOS	KP 12 KP 12 62200 KAUHAVA

Kuvio 1: Esimerkki tallennusselvityksen "virhelistasta"

Lyhennyssääntöjen soveltamisen lisäksi olivat erot, jotka johtuivat välimerkin vaihtelevasta käytöstä suhteellisen yleisiä (vrt. 'lasten hoitaja' 'lastenhoitaja' 'lasten-hoitaja'). Eroja tallennuskenttien välillä saattoi myös syntyä siitä, että lomakkeelle tallennettiin turhia tietoja (esim. työpaikan kunta kysymyksessä 6). Joitakin selviä lyöntivirheitä löytyi myös alkuperäisestä tallennustiedostosta. Vaikutelmaksi jäi, että ruotsinkielisten lomakkeiden kohdalla esiintyi enemmän kirjoitusvirheitä kuin suomenkielisten lomakkeiden

kodalla (esim. 'jordrukare' 'lektori i vårdlära'). Kaiken kaikkiaan selviä lyöntivirheitä esiintyi kuitenkin verraten vähän.

Lopputoteamuksena voidaan sanoa, että poikkeavia tallennuskenttiä esiintyi melko paljon varsinkin aakkostietojen kohdalla. Myös numeeristen tietojen kohdalla poikkeavuuksien osuutta - 3.2 % - voidaan pitää melko korkeana. Mikäli myös tyhjät kentät otetaan huomioon, poikkeavien numeeristen tietojen osuus laskee 0.6 %:iin. Tallennetut tiedot muodostivat kuitenkin vain sen raaka-aineen, josta lopulliset laskentamuuttujat jalostettiin. Esim. aakkostietojen kohdalla tallennusvirheet saattoivat jäädä kokonaan vaille vaikutusta tai johtivat vain ko. tietueen manuaalikäsittelyyn.

Tulokset osoittavat myös selvästi, ettei lomakkeiden manuaalista tallentamista voida pitää pelkkänä mekaanisena toimintana, jota voitaisiin verrata vaikkapa optiseen lukemiseen.

Tallentajan tiedot esim. yritysten ja paikkakuntien nimistä antavat välttämätöntä perustietoa ja "esiymmärrystä", joka on tärkeä silloin, kun lomakkeelle kirjoitettu selväkielinen teksti on luettava lyhyessä ajassa ja tiedot on tallennettava nopeasti. Mitä enemmän tietoja tallentajilla oli yhteiskunnastamme, sitä luotettavammin tiedot varmaankin tallennettiin. Yllättävän moni vastaaja kirjoitti esim. työpaikkansa nimen ja osoitteen virheellisesti. Tallentajan tiedot paikannimistä saattoivat joskus myös johtaa harhaan. Paikkakunta 'Kulennoinen' (kirjoitettu epäselvällä käsialalla) tuotti esim. tallentajille vaikeuksia - ensimmäisellä yrityksellä tallennettiin 'Kuhmoinen'.

Seuraavassa luvussa uusia muuttujia verrataan vastaaviin alkuperäisiin tietoihin.

## 5. MUUTTUJAKOHTAINEN VERTAILU

### *Vertailun periaatteet ja tunnuslukujen laskeminen*

Tietojen vertailu tapahtui siten, että jokainen muuttuja taulukoitiin ristiin "itsensä kanssa". Kun kyseessä olivat tiedot pääasiallisesta toiminnasta, työssäolo- ja työttömyyskuukausista tai ammatissa-toimivuudesta, mukana oli tutkimuksen koko perusjoukko (N = 2 188). Muiden tietojen kohdalla perusjoukko rajattiin siten, että vain ne henkilöt, jotka vähintään toisessa tiedostossa oli luokiteltu työlliseksi otettiin mukaan (N = 1 461). Joitakin osajoukkoja on myös tutkittu erikseen.

Tilan säästämiseksi näitä ristiintaulukoita esitetään tämän tutkimuksen liiteosassa. Sen sijaan jokaisesta tiedosta on esitetty taulukko, josta selviää molempien muuttujien suorat jakaumat ja samalla tavalla luokiteltujen tapausten osuus. Jokaiselle muuttuja-arvolle on myös laskettu ns. suhteellinen nettovirhe.

Samalla tavalla luokiteltujen osuus on laskettu siten, että saman koodiarvon saaneiden määrä on kerrottu kahdella ja jaettu niiden tapausten määrällä, joilla joko varsinaisessa tai uusintatutkimuksen aineistossa oli kyseinen koodi. Suhteellista nettovirhettä laskiessa on tietyn muuttujaluokkaan sijoitettujen henkilöiden määrän erotus jaettu alkuperäisen aineiston ko. luokan määrällä. Nettovirhe on positiivinen jos varsinaisen aineiston luvut ovat suurempia kuin uusintatutkimuksen luvut, kun taas päinvastaisessa tapauksessa se on negatiivinen.

Samalla tavalla luokiteltujen tapausten osuus kuvaa lähinnä muuttujan reliabiliteettia, kun taas nettovirhe kertoo kuinka paljon eri muuttujaluokkien kokonaismäärät poikkeavat toisistaan. Suuret nettovirheet ovat osoituksena siitä, että poikkeavuudet tiedostojen välillä eivät johdu pelkästään sattumasta. Suhteellinen bruttovirhe saadaan, kun samalla tavalla luokiteltujen osuus vähennetään luvusta 100. Koska uusintatutkimuksenkin tiedoissa saattaa esiintyä virheitä olisi ehkä oikeampaa puhua netto- ja bruttoeroista. Tässä selvityksessä on kuitenkin seurattu jo vakiintunutta kielenkäyttöä.

Seuraavassa tarkastellaan laskennan 'lomakemuuttujia' yksi kerrallaan.

### Pääasiallinen toiminta

Pääasiallisen toiminnan käsite kuvaa henkilön taloudellisen toiminnan laatua. Tiedot pääasiallisesta toiminnasta perustuvat ensi sijaisesti laskentalomakkeen kysymykseen nro 2.

Tallennusselvityksen mukaan identtisten TOI-kenttien osuus (kysymys 2) oli 99.0 %. Oli siis syytä olettaa, että uusintatutkimuksen tieto pääasiallisesta toiminnasta olisi lähes identtinen alkuperäisen tiedon kanssa. Vertailun tulos selviää taulukosta 3.

### Taulukko 3: Pääasiallinen toiminta

Pääasiallinen toiminta (15-74-v.)	Vanha tieto	Uusi tieto	Sama luok. %	Nettovirhe %
Työvoima	1541	1540	99.6	0.1
Työllinen työvoima	1455	1454	99.6	0.1
Työtön työvoima	86	86	98.8	0.0
Työvoimaan kuulumattomat	645	648	99.0	-0.5
Opiskelijat, koululaiset	213	212	99.8	0.5
Kotitaloustyötä tekevät	61	63	93.6	-3.3
Eläkeläiset	329	328	99.2	0.3
Varusmiehet	20	20	100.0	0.0
Muut työvoiman ulkopuolella olevat	22	25	89.4	-13.6
Tuntematon	2	0	..	..
<b>Yhteensä</b>	<b>2188</b>	<b>2188</b>	<b>99.2</b>	<b>-</b>

Taulukosta 3 nähdään, että 99.2 % tapauksista sai saman muuttuja-arvon molemmilla käsittelykerroilla. Mikäli käytetään karkeata luokitusta 'työvoima', 'työvoimaan kuulumattomat' ja 'tuntematon', samaan luokkaan kuuluvien määrä on 99.3 %. Mikäli käytämme vielä karkeampaa luokitusta 'työllinen työvoima' - 'muut' prosenttiosuus on 99.4 %.

Samalla tavalla luokiteltujen osuus on alhaisin luokissa 'kotitaloustyötä tekevät' ja 'muut työvoiman ulkopuolella olevat'. Kooditus olisi ehkä ollut luotettavampi, mikäli myös mahdollinen lisäinformaatio lomakkeen kohdassa '0 Muu, mikä?' olisi tallennettu. Maatalouden emännille ei aina tahtonut löytyä yksiselitteistä koodia.

Mikäli tarkastellaan yksityiskohtaisinta luokitusta, poikkeavia tapauksia oli otoksessa yhteensä 18 (noin 1 % koko väestöstä). Valtaosa poikkeavuuksis-

ta johtui mitä todennäköisimmin poikkeavista tallennustiedoista - eikä esim. myöhemmin tehdyistä tarkistuksista.

Henkilöille, joiden ikä laskenta-ajankohtana oli alle 15 vuotta tai yli 74 vuotta ei lähetetty laskentalomaketta. Heidän "pääasiallinen toimintansa" on kuitenkin määritelty (lapsi/eläkeläinen). Mikäli lapset ja yli 74-vuotiaat otetaan mukaan tarkasteluun, pääasiallisen toiminta-tiedon reliabiliteetti on luonnollisesti vielä korkeampi. Sama koskee myös työ- ja työttömyyskuukausi- sekä ammattissatoimivuustietoja, joita tarkastellaan seuraavissa jaksoissa.

#### *Työssäolo- ja työttömyyskuukaudet*

Tallennusselvityksen mukaan identtisten KTYÖ-, OSTYÖ-, OSA-, TYKE- ja METS-kenttien osuudet (lomakkeen kysymykset 8 ja 4.1) vaihtelivat 98.9 ja 99.7 prosentin välillä, kun myös tyhjä tallennuskentät otetaan huomioon.

#### **Taulukko 4: Työssäolo- ja työttömyyskuukaudet**

Työssäolo- ja työttömyyskuukaudet (15-74-v.)	Vanha tieto	Uusi tieto	Sama luok. %	Nettovirhe %
Kokopäivätyö	2188	2188	99.6	-
Osa-aikatyö (väh. 20 h/v)	2188	2188	99.6	-
Osa-aikatyö (alle 20 h/v)	2188	2188	99.7	-
Työttömyyskuukaudet	2188	2188	99.7	-
Metsätyökuukaudet	1461	1461	99.0	-

Taulukosta 4 voidaan nähdä, että samalla tavalla luokiteltujen tapausten osuus on erittäin korkea ja bruttovirhe siten lähes olematon. Kuukausittaisia nettovirheitä ei ole tulostettu erikseen. Käytännössä nettovirheitä ei juurikaan esiintynyt.

Muuttujaa 'osa-aikatyö (alle 20 h/v)' ei ole käytetty väestölaskennan taulukkotuotannossa. Tieto metsätyökuukausista oli luonteeltaan apumuuttuja, jota käytettiin lähinnä elinkeinotiedon päättelyssä.

### Ammatissatoimivuus

Tieto ammatissatoimivuudesta perustui lähinnä työ- ja työttömyyskuukausien yhteenlaskettuun määrään. Ammatissa toimivaan väestöön luettiin kaikki 15-74-vuotiaat, jotka olivat yhteensä vähintään kuusi kuukautta työssä (työaika vähintään 20 h/viikko) ja/tai työttömänä.

Joitakin muita sääntöjä käytettiin myös. Esimerkiksi kaikki ne työlliset, jotka eivät olleet vastanneet työ- ja työttömyyskuukausia koskeviin kysymyksiin, mutta jotka olivat hyväksyneet lomakkeelle esitetyt työpaikkatiedot vuodelta 1980 luokiteltiin ammatissa toimiviksi.

#### Taulukko 5: Ammatissatoimivuus

Ammatissatoimivuus (15-74-vuotiaat)	Vanha tieto	Uusi tieto	Sama luok. %	Netto- virhe %
Ammatissa toimiva	1422	1423	99.8	-0.1
Ammatissa toimimaton	759	758	99.7	0.1
Tuntematon	7	7	..	..
Yhteensä	2188	2188	99.7	-

99.7 % kaikista havainnoista luokiteltiin molemmilla kerroilla samaan ryhmään. Poikkeavista tapauksista (N=7) kolme henkilöä sai eri koodin poikkeavien työssäolo- ja/tai työttömyyskuukausien vuoksi. Yksi henkilö oli luokiteltu ammatissa toimivaksi rekisteritäydennyksen yhteydessä. Syytä miksi loput kolme henkilöä saivat eri koodin ei ole voitu yksityiskohtia myöten selvittää.

Taulukon 5 luvuista nähdään, että muuttujaluokkien brutto- ja nettovirheet ovat erittäin pienet.

## Ammattiasema

Ammattiasemaluokitus kuvaa työllisten asemaa työelämässä. Tieto ammattiasemasta perustuu lähinnä lomakekysymykseen nro 2. Tallennusselvityksen mukaan identtisten TOI-kenttien osuus oli 99.0 %.

**Taulukko 6: Ammattiasema**

Ammattiasema	Vanha tieto	Uusi tieto	Sama luok. %	Nettovirhe %
Palkansaaja	1242	1238	99.6	0.3
Yksinäisyrittäjä	118	116	96.6	1.7
Yrittäjä, jolla palkattua työvoimaa	43	49	91.3	-14.0
Yrittäjän avustava perheenjäsen	52	51	89.3	1.9
Tuntematon	-	-	..	..
Ei työllinen	6	7	(0)	..
<b>Yhteensä</b>	<b>1461</b>	<b>1461</b>	<b>98.3</b>	<b>-</b>

98.3 % tapauksista sai molemmilla käsittelykerroilla saman koodiarvon. Poikkeavia tapauksia oli yhteensä 25. Mikäli tarkastellaan vain niitä henkilöitä, jotka molemmissa aineistoissa oli luokiteltu työllisiksi, poikkeavien tapauksien määrä supistuu 12:een (noin 1 % työllisistä).

Ammattiasematietoa korjailtiin eri tarkistussääntöjen mukaan. Lopputarkistuksessa käytettiin esim. seuraavaa päättelysääntöä:

Jos useammalla samaan asuntokuntaan kuuluvista on sama toimialakoodi, tarkistetaan, että vain yksi on yrittäjä tai työnantajayrittäjä ja muut avustavia perheenjäseniä.

Useasta yrittäjästä valitaan varsinaiseksi yrittäjäksi vanhin mies. Kuitenkin alle 25-vuotiasta ei valita yrittäjäksi, jos asuntokunnasta löytyy vanhempia naishenkilöitä.

Toisen säännön mukaan osakeyhtiössä työssä olleet luokiteltiin palkansaajiksi lomakkeen ammattiasema-tiedosta riippumatta.

Mikäli kyseisiä korjausajoja ei olisi tehty uusintatutkimuksen (ja varsinaisen käsittelyn) yhteydessä, yrittäjien avustavien perheenjäsenten määrä olisi ollut noin kolmanneksen pienempi.

Taulukon 6 mukaan suurin nettovirhe (-14.0 %) on luokassa 'Yrittäjä, jolla palkattua työvoimaa'. Tulosta ei kuitenkaan voida tulkita niin, että kyseinen muuttujaluokka alkuperäisessä aineistossa

on selvästi liian pieni - uusintatutkimuksen tuloksiin ei voi täysin luottaa. Tulos kertoo kuitenkin, että kyseisen ryhmän kokoa on vaikeata laskea. Kun muuttujan koodiarvo on 'lähtötiedon' lisäksi riippuvainen myös muista tiedoista - tässä tapauksessa muiden perheenjäsenten ammattiasemasta ja oman työpaikan juridisesta muodosta, on mahdollista että uusintatutkimus tuottaa poikkeavan tuloksen vaikkakin lähtötieto olisi sama. Tähän ongelmaan palataan vielä tuonnempana.

### *Ammatti*

Ammattitiedot perustuvat sekä lomake- että rekisteritietoihin. Väestön keskusrekisteristä saatiin henkikirjoitukseen perustuva tieto henkilön ammatinimikkeestä 1.1.1985. Jos tämä nimike voitiin koneellisesti koodittaa yksiselitteisesti, se esitäytettiin laskentalomakkeelle. Henkilön tuli ilmoittaa ammattitietonsa vain, jos esitäytetyt tiedot puuttuivat tai olivat virheelliset.

Silloin kun työlliseksi luokiteltu henkilö ei ollut ilmoittanut ammattiansa eikä lomakkeelle ollut esitäytetty ammattia, saatettiin ammatiksi merkitä vuoden 1980 väestölaskennan ammattikoodi. Joissakin tapauksissa ammattitieto poimittiin Väestörekisterikeskuksen rekisteristä tai pääteltiin toimiala-, tutkinto- ym. tietojen perusteella.

Selväkielisen ammattinimikkeen kohdalla identtisten tallennuskenttien osuus oli 68.6 %. Taulukon 7 luvuista nähdään saman ammattikoodin saaneiden osuus.

## Taulukko 7: Ammatti

Ammatti	Vanha tieto	Uusi tieto	Sama luok. %	Nettovirhe %
1-numerotaso	1461	1461	94.1	-
0 Tekninen työ	311	307	94.2	1.3
1 Hallinnollinen työ	211	219	92.0	-3.8
2 Kaupallinen työ	105	100	91.7	4.8
3 Maa- ja metsätaloustyö, kalastusala	143	148	96.2	-0.3
4 Kaivos-, syväkairaus- ja rikastustyö	4	5	..	..
5 Kuljetus- ja liikennet.	93	95	96.8	-2.2
6/7 Teollinen työ, koneenhoito ym.	419	411	96.6	1.9
8 Palvelutyö	162	158	95.0	2.5
9 Muualla luokittelematon työ ja työvoima	7	11	..	..
Ei työllinen	6	7	(0)	..
2-numerotaso	1461	1461	91.7	-
3-numerotaso	1461	1461	88.7	-
4-numerotaso	1461	1461	88.4	-

Samalla tavalla luokiteltujen osuus on alhaisin ryhmissä 'hallinnollinen työ' ja 'kaupallinen työ'. Koodit olivat sekoittuneet osittain keskenään, osittain taas ryhmään 'tekninen työ'.

Mitä tarkemmalle luokitustasolle mennään, sitä pienempi on niiden tapausten osuus, jotka molemmilla käsittelykerroilla saivat saman koodiarvon. Mikäli koodi poikkeaa 1-numerotasolla, se tietenkin poikkeaa myös tarkemmilla luokitustasoilla. 4-numerotasolla poikkeavia oli yhteensä 169 tapausta (noin 12 % työllisistä).

Mikäli perusjoukko rajataan siten, että mukana on vain työllisiä, samalla tavalla luokiteltujen osuus on 1-numerotasolla 94.3 % ja 4-numerotasolla 89.2 %.

Seuraavassa esitetään joitakin satunnaisesti poimittuja esimerkkejä poikkeavista ammattikodeista. Ylimpänä on joko lomakkeelta tallennettu vastausvelvollisen kirjoittama ammatti tai vastausvelvollisen hyväksymä esitäytetty ammattinimike (näitä ei ole voitu erottaa toisistaan). Sen alla on koodi, jonka henkilö sai varsinaisessa käsittelyssä kun taas alimpana on uusintatutkimuksen ammattikoodi. Oikeanpuoleisessa sarakkeessa on vastaavat toimialatiedot.

Ammattitieto kooditettiin vasta sen jälkeen, kun henkilölle oli päätelty toimiala. Sekä automaattisessa että manuaalisessa ammattikoodituksessa toimialaa käytettiin hyväksi aputietona.

Ammatti	Toimiala
<u>MATERIAALITOIMINTOJEN ESIMIES</u>	
001 Rakennusinsinöörit	35402 Kattohuovan valmistus
563 Tieliikenteen johto	35402 "
<u>TUTKIMUS-KEHITYSPÄÄL</u>	
013 Koneteknikot	38242 Rakennus- ja kiven jalostuskoneiden ja -laitteiden valmistus
111 Tekniset johtohenkilöt	38242 "
<u>LABORATORIOHOITAJA</u>	
028 Laborantit ja laboratorioapulaiset	93312 Erikoislääkäreiden palvelukset
0323 Laboratoriohoitajat	93312 "
<u>OHJAAJA</u>	
0913 Muut sosiaalialan työntekijät	93431 Päihdyttävien aineiden väärinkäyttäjien laitoshoido
0912 Sosiaalialan erityistyöntekijät	93431 "
<u>KASSÖRSKA</u>	
121 Pankkien, postien ja konttoreiden kassanhoitajat	38299 Muu koneiden valmistus
122 Myymälöiden ja ravintoloiden kassanhoitajat	381 Metallituotteiden valmistus
<u>KONTTORISTI</u>	
151 Yleiskonttoristit	3559 Muu kumituotteiden valmistus
159 Muut konttorityön ammatit	3559 "
<u>LIHAMESTARI</u>	
201 Vähittäiskauppiat	31112 Lihanjalostus
231 Myymälähenkilöstö	31112 "
<u>OSASTOPÄÄL</u>	
232 Myymälänhoitajat ja osastonhoitajat	62112 Tavaratalot
119 Muut liikeyritysten ja järjestöjen hallinnollisen työn ammatit	62511 Varsinainen rautakauppa
<u>OMPELIJA VIIMEISTELIJÄ</u>	
615 Valmisvaateteollisuuden ompelijat	32204 Tekstiilipäällysvaatteiden valmistus
605 Viimeistämön ja värjäämön työntek.	32204 "
<u>Ei ammattinimikettä</u>	
653 Ohut- ja paksulevysepät	7131 Ilmakuljetus
655 Hitsaajat ja kaasuleikkaajat	7131 "
<u>HITSAUSKYPÄRÄN VALMISTAJA</u>	
659 Muut konepaja- ja rakennusmetallityön ammatit	35609 Muu muovituotteiden valmistus
752 Muovituotetyöntekijät	35609 "
<u>PORARI</u>	
677 Konepusepät ym.	5129 Muut sivu- ja alaurakat
400 Kaivosmiehet, panostajat ym.	5129 "
<u>LEIKKAAMON TYÖNJOHT</u>	
710 Lasihyttityöntekijät	36201 Tasolasin valmistus ja jalostus
696 Lasinasettajat	36201 "
<u>TYÖKALUJEN JAKAJA</u>	
781 Varastotyöntekijät	2302 Muiden malmien kaivuu
581 Lehdenjakajat	2302 "
<u>HUSMOR</u>	
9999 Ei työllinen	11 Maatalous ja metsästys
310 Maanviljelystyöntekijät ja kotieläintenhoitajat	11 "

Edellisellä sivulla olevista esimerkeistä on käytettävissä olevien tietojen avulla usein mahdollonta sanoa, kumpi ammattikoodeista on, 'oikea', kumpi taas 'väärä'. Itse asiassa koko otoksesta löytyi vain muutama tapaus, jonka kohdalla ammattikoodi oli selvästi virheellinen. Ammattien koodituksessa ongelmana näyttääkin olevan se, että itse ammattiluokitus jaottelee ammatteja tarkemmin kuin mitä käytettävissä olevien tietojen perusteella on mahdollista tehdä.

Seuraavassa esitetään vielä esimerkkejä ammateista, jotka usein saivat poikkeavan koodin uusintatutkimuksen yhteydessä. Käsittelyvirheselvityksen pienestä otoskoosta johtuen luettelosta ei kuitenkaan voitane vetää pitkälle meneviä päätöksiä.

Ammattikoodi ja -nimike	(1) N	(2) %	(3) %
151 Yleiskonttoristit	8	4.7	3.4
0913 Muut sosiaalialan työntekijät	6	3.6	1.1
201 Vähittäiskauppiat	6	3.6	1.2
310 Maanviljelystyöntekijät ja kotieläintenhoitajat	6	3.6	3.1
659 Muut konepaja- ja rakennusmetallityön ammatit	6	3.6	1.0

(1) ammattikoodin esiintymistiheys 'virhelistalla'

(2) ammattikoodin osuus poikkeavien tapausten määrästä

(3) ammattikoodin osuus koko aineistosta

Salme Kiiski (1988), joka on tutkinut manuaalisen ammattikoodituksen tasoa työvoimatutkimuksen yhteydessä sai tulokseksi, että 2 % ammattikoodeista oli virheellisiä 1-numerotasolla. 2-numerotasolla osuus oli 3.2 %, 3-numerotasolla 3.9 % ja 4-numerotasolla 4.0 %. Virheitä oli eniten teollisen ja rakennustyön ammateissa (6/7). Niissä edellytetään Kiisken mukaan usein työpaikan toimialan huomioonottamista, koska itse ammattinimike on usein liian 'yleisellä' tasolla, esim. metallimies, harjoittelija, rakennusmies, ompelija, huoltomies jne.

Näyttää siltä, että työvoimatutkimuksessa ammattien kooditus on johdonmukaisempaa kuin väestölaskennoissa. Tämä johtunee sekä siitä, että haastattelututkimuksessa on mahdollista saada täsmällisempi alkutieto kuin postikyselyssä sekä siitä, että työvoimatutkimuksen koodajilla on pitkäaikainen työkokemus - työvoimatutkimushan tehdään joka kuukausi.

Työvoimatutkimuksen käsittelyvirheselvityksen tulokset osoittavat myös, etteivät ammattikoodien bruttovirheet johdu pelkästään itse luokituksen heikkouksista - parempaan tulokseen olisi voitu päästä myös väestölaskennassa.

#### *Elinkeino (toimiala)*

Väestölaskennassa omaksutun määritelmän mukaan saavat kaikki samassa toimipaikassa työskentelevät saman elinkeinon (toimialan) ammatista riippumatta. Tieto elinkeinosta perustuu lähinnä laskentalomakkeen työpaikkatietoihin. Henkilön tuli ilmoittaa työpaikkatietonsa vain, jos esitetyt tiedot puuttuivat tai olivat virheellisiä.

Selväkielisen toimialatiedon kohdalla identtisten tallennuskenttien osuus oli 74.0 %. Taulukosta 8 voidaan nähdä, että saman elinkeinon saaneiden osuus oli onneksi korkeampi.

#### **Taulukko 8: Elinkeino (toimiala)**

Elinkeino (toimiala)	Vanha tieto	Uusi tieto	Sama luok. %	Nettovirhe %
1-numerotaso	1461	1461	95.8	-
1 Maatalous	146	149	95.6	-2.1
2 Kaivostoiminta	8	8	..	..
3 Teollisuus	370	369	97.7	0.2
4 Sähkö- yms. huolto	27	27	92.6	0.0
5 Rakennustoiminta	104	99	92.6	4.8
6 Kauppa	191	183	95.2	4.2
7 Kuljetus	105	106	97.6	-0.9
8 Rahoitus	101	101	95.0	0.0
9 Yhteisk. palvelut	404	412	96.6	-2.0
Ei työllinen	6	7	(0)	..
2-numerotaso	1461	1461	93.7	-
3-numerotaso	1461	1461	91.8	-
4-numerotaso	1461	1461	89.6	-
5-numerotaso	1461	1461	85.9	-

Samalla tavalla luokiteltujen osuus on pienin toimialaluokissa 4 (sähkö- yms. huolto) ja 5 (rakennustoiminta). Nettovirhe taas on korkein toimialaluokissa 'rakennustoiminta' ja 'kauppa'. Selvää systemaattisuutta virheiden suhteen ei ollut havaittavissa. Liitetaulukosta 5 voidaan kuitenkin nähdä, että meko suuri osa toimialaluokkien 5 (rakennustoiminta) ja 6 (kauppa) poikkeavista tapauksista uusintatutkimuksen yhteydessä luokiteltiin toimialaluokkaan 9 (palvelut).

Samoin kuin ammattitiedon kohdalla, poikkeavien tapausten määrä kasvaa mitä tarkemmalle luokitustasolle mennään. 5-numerotasolla poikkeavia tapauksia oli otoksessa yhteensä 206 (noin 14 % työllisistä).

Mikäli perusjoukkoa rajataan siten, että mukana on vain työllisiä, samalla tavalla luokiteltujen osuus nousee noin yhdellä prosenttiyksiköllä: 1-numerotasolla se on 96.4 % ja 5-numerotasolla 86.5 %.

Seuraavassa esitetään esimerkkejä poikkeavista elinkeinokodeista. Ylin toimialatieto on aina henkilön itse ilmoittama toimiala. Oikeanpuolisessa sarakkeessa on vastaavat ammattitiedot.

Toimiala	Ammatti
<u>Ei toimialatietoa</u>	
3 Teollisuus	582 Toimistovahtimestarit ja -lähetit
34112 Paperin ja kartongin valmistus	760 Pakkaajat ja paketoijat ym.
<u>KAIVINKONEURAKOINTI</u>	
522 Louhinta ja maansiirto	151 Yleiskonttoristit
99999 Ei työllinen	9999 Ei työllinen
<u>KIRJAPAINO</u>	
3420 Graafinen tuotanto, kustannustoim.	151 Yleiskonttoristit
34201 Painaminen	159 Muut konttorityön ammatit
<u>PYÖRÄTEHDAS</u>	
3844 Polku- ja moottoripyörien valmistus	657 Kokoojat ja vaihetyöntekijät
3903 Urheiluvälineiden valmistus	657 "
<u>POSTI</u>	
72000 Tietoliikenne	580 Postinkantajat ja -lajittelijat
72001 Postiliikenne	580 "
<u>SÄHKÖLIIKE</u>	
9512 Kotitaloussähkölaitteiden korj.	660 Sähköasentajat
5122 Sähköasennustyöt	660 "
<u>KIRJAPAINON MYYNTITOIMISTO</u>	
34204 Kustannustoiminta	130 Sihteerit
342 Graafinen tuotanto, kustannustoiminta	130 "
<u>Ei toimialatietoa</u>	
38319 Muiden teollisuussähkölaitteiden valmistus	657 Kokoojat ja vaihetyöntekijät
38311 Sähkökoneiden ja muuntajien valm.	657 "
<u>VAATEMYYMÄLÄ</u>	
62431 Miesten ja naisten pukukauppa	231 Myymälähenkilöstö
6241 Tekstiili- ja vaatetusalan yleisvähittäiskauppa	231 "
<u>Ei toimialatietoa</u>	
72002 Teleliikenne	664 puhelin- ja linja-asentajat
4101 Sähkön tuotanto ja jakelu laitoshoido	664 "
<u>Ei toimialatietoa</u>	
93319 Muu terveyden- ja avosairaanhoido	0332 Vastaanottoapulaiset
93312 Erikoislääkäreiden palvelukset	0332 "
<u>KONEPAJA</u>	
3821 Kiinteiden moottorien ja turbiinien valmistus	655 Hitsaajat
381 Metallituotteiden valmistus	655 "
<u>KORKEAKOULU</u>	
9318 Korkean asteen oppilaitokset	028 Laborantit
94202 Museot, eläin ja kasvitarhat	0992 Muut humanistisen, luonnontieteellisen ja yhteiskuntatieteellisen työn tekijät
<u>AMMATTILEHTI</u>	
93502 Muut elinkeino- ja ammattijärjestöt	0841 Sanomalehden toimittajat
34202 Kuvalaattojen valmistus	0841 "

Aivan kuten ammattitietojen kohdalla, tulkinanvaraisia tapauksia on paljon. Määrä ei välttämättä pienene, vaikka työpaikan nimi- ja osoitetiedot otetaan avuksi.

Seuraavassa esitetään vielä ne toimialat, jotka useimmin esiintyvät 'virhelistalla'.

Koodi	Elinkeino (toimiala)	(1) N	(2) %	(3) %
511	Varsinaiset talonrakennustyöt	7	3.4	3.7
3710	Raudan, teräksen ja ferroseosten valmistus	5	2.4	0.3
38299	Muu koneiden valmistus	5	2.4	1.2
93311	Sairaalapalvelu	5	2.4	3.5

- (1) toimialakoodin esiintymistiheys 'virhelistalla'  
 (2) toimialan osuus poikkeavien tapausten määrästä  
 (3) toimialan osuus koko aineistosta

Salme Kiisken tutkimuksen mukaan työvoimatilastojen luotettavuudesta virheellisiä toimialakoodeja 1-numerotasolla oli 1.8 %. 2-numerotasolla osuus oli 3.3 % ja 3-numerotasolla 4.3 %. Myös toimialatiedon kohdalla vaikuttaa siltä, että työvoimatutkimuksessa käsittelyvirheiden osuus on pienempi kuin väestölaskennoissa.

Esimerkkeinä tulkinallisesti vaikeista tapauksista Kiiski mainitsee julkisen toiminnan toimipaikat, joissa oli vaikeata päätellä, milloin oli kysymys varsinaisesta hallinnosta (911) tai muusta toiminnasta, esim. rakentamisesta (510, 520). Lääkintäpalvelujen (933) ja sosiaalihuollon raja osoittautui myös epäselväksi.

Tallennusselvityksen yhteydessä todettiin, että ruotsinkieliset lomakkeet ilmeisesti tuottivat tallentajille ongelmia. Näyttää siltä, että ruotsinkieliset nimikkeet olivat hankalia myös jatkokäsittelyn yhteydessä. Taulukossa 9 on samalla tavalla luokiteltuja ammatti- ja toimialatietoja taulukoitu kielen mukaan.

**Taulukko 9: Ammatti- ja toimialatiedot kielen mukaan**

Luokitustaso, kieli	Ammatti Sama luok. %	Toimiala Sama luok. %	Henkilöitä yhteensä
1-numerotaso suom.	93.8	95.9	1361
ruots.	88.8	93.9	98
2-numerotaso suom.	92.1	93.8	1361
ruots.	86.7	91.8	98
3-numerotaso suom.	89.1	91.4	1361
ruots.	82.7	89.8	98
4-numerotaso suom.	88.8	89.1	1361
ruots.	82.7	89.8	98

suom. = suomenkieliset työlliset  
ruots. = ruotsinkieliset työlliset

Taulukosta nähdään, että ruotsinkielisen ammatti- ja elinkeinotiedon laatu on huonompi kuin suomenkielisen vastaavan tiedon. Ammattitiedon kohdalla ero on selvempi kuin toimialatiedon kohdalla.

*Ns. liikkuvuustyyppi*

Ns. liikkuvuustyyppi määritellään työn luonteen ja suorittamispaikan mukaan. Tiedot liikkuvuustyyppistä perustuvat laskentalomakkeella saatuihin tietoihin henkilön työpaikasta.

**Taulukko 10: Liikkuvuustyyppi**

Liikkuvuus- tyyppi	Vanha tieto	Uusi tieto	Sama luok. %	Netto- virhe %
Työpaikka asunnon yhteydessä	141	161	86.1	-14.2
Yhdessä kunnassa liikkuva	92	76	65.4	17.4
Kunnassa tuntematon	16	3	..	..
Useissa kunnissa liikkuva	15	19	..	..
Kiinteä työpaikka, mutta ei toimipaikka	33	47	70.0	-42.4
Ulkomailla	3	3	..	..
Toimipaikka	1150	1129	97.1	1.8
Tuntematon	5	16	..	..
Ei työllinen	6	7	(0)	..
<b>Yhteensä</b>	<b>1461</b>	<b>1461</b>	<b>91.6</b>	<b>-</b>

Poikkeavia tapauksia oli yhteensä 125 (noin 9 % työllisistä). Mikäli luokitusta supistetaan siten, että luokka 'työpaikka asunnon yhteydessä' yhdistetään luokkaan 'toimipaikka' ja luokka 'kunnassa

tuntematon' luokkaan 'kiinteä työpaikka, mutta ei toimipaikka' samalla tavalla luokiteltujen osuus on 92.8 %. Selviä virhemerkintöjä havaittiin sekä varsinaisessa että uusintatutkimuksen aineistossa. Verraten paljon rajatapauksia esiintyi myös. Mikäli uusintatutkimuksen ilmiselvät virheet korjataan, poikkeavia tapauksia oli noin 100 (noin 7 % työllisistä).

Poikkeavista tapauksista noin kolmannes oli sellaisia, että myös toimiala oli kooditettu eri tavalla. Joka neljännellä oli taas poikkeava ammattikoodi. Yhteensä 70 henkilöllä (5 % työllisistä) oli poikkeava liikkuvuustyyppikoodi vaikkakin ammatti- ja toimialatiedot olivat identtiset.

Seuraavassa luetellaan joitakin poikkeavuustapauksia. Jokaisesta tapauksesta on lisäksi muutama esimerkki.

- *Työpaikka asunnon yhteydessä / toimipaikka*

Esim. (1) Työpaikka postiasema, ammatti konttorinhoitaja. (2) Elinkeino kalanviljely, ammatti kalastusmestari.

- *Yhdessä kunnassa liikkuva / toimipaikka*

Esim. (1) Työpaikka rannikkolaivasto, ammatti alempi toimiupseeri (2) Työpaikka Helsingin yliopisto, ammatti opettaja, assistentti.

- *Yhdessä kunnassa liikkuva / kiinteä työpaikka, mutta ei toimipaikka*

Esim. (1) Toimiala 'varsinaiset rakennustyöt', ammatti rakennuspuutyöntekijä. (2) Toimiala 'rakennusten puhdistus ja siivous' ammattiryhmä 'siivoojat ym.'.

- *Työpaikka asunnon yhteydessä / yhdessä kunnassa liikkuva*

Esim. (1) Maanviljelijä. (2) Toimiala maalaus-työt, ammatti maalari.

Liikkuvuustyyppi -tiedon ongelmana on mm. se, että muuttuja sisältää kaksi eri luokitusta:

1. Työpaikan sijainti (asunnon yhteydessä, ei asunnon yhteydessä, liikkuva)
2. Työpaikkatyyppi (toimipaikka, kiinteä työpaikka mutta ei toimipaikka, ei kiinteä työpaikka).

Mikäli henkilö "asuu toimipaikassa" löytyykin kaksi eri koodivaihtoehtoa: 'työpaikka asunnon yhteydessä' ja 'toimipaikka'. Mikäli työpaikka on löytynyt toimipaikkatiedostosta, on ilmeisesti useimmiten valittu jälkimmäinen vaihtoehto. Edellä olevat esimerkit viittaavat kuitenkin siihen, ettei tällaisia tapauksia aina ole voitu koodittaa yksiselitteisesti.

### *Työpaikan sijaintikunta*

Työpaikan sijaintikunta -tieto perustuu laskentalomakkeen työpaikkaa koskeviin kysymyksiin. Sijaintikuntatiedon avulla saadaan selville tietyllä alueella työssäkävien määrä tai paljonko kunnassa on - kuten on ollut tapana sanoa - 'työpaikkoja'.

Mikäli työpaikalle ei voitu määritellä sijaintikuntaa, tarjolla oli vielä seuraavat vaihtoehdot:

- Ruotsi
- Norja
- Neuvostoliitto
- muut ulkomaat
- useissa kunnissa liikkuva
- tuntematon

Taulukosta 11 selviää, kuinka usein uusintatutkimuksen tieto työpaikan sijainnista poikkesi alkuperäisestä tiedosta.

### **Taulukko 11: Työpaikan sijaintikunta**

Työpaikan sijainti	Vanha tieto	Uusi tieto	Sama luok. %	Nettovirhe %
Kaikki työlliset	1461	1461	94.4	-
Vain ne, joilla kuntakoodi	1364	1364	97.2	-
Vain toimipaikat	1075	1075	97.1	-

Poikkeavia tapauksia oli yhteensä 82 (noin 6 % työllisistä). Jos useissa kunnissa liikkuvat, ulkomailla olevat sekä muut työpaikat, joille ei voitu määritellä sijaintikuntaa poistetaan poikkeavia oli yhteensä 40 (noin 3 % tästä perusjoukosta). Mikäli tarkastellaan vain toimipaikkoja poikkeavien tapausten määrä supistuu 32:een (noin 3 %).

Syytä siihen, miksi työpaikkojen sijaintikunta uusintatutkimuksessa erosi alkuperäisestä tiedosta ei ole tarkemmin selvitetty. Myös uusintatutkimuksen tiedoissa saattaa olla virheitä. Varsinkin toimipaikkojen kohdalla poikkeavien tapausten osuus on yllättävän korkea.

### *Työpaikan juridinen muoto*

Työpaikan juridisen muodon luokitus kuvaa työpaikan omistajuutta ja yritysmuotoa. Juridisen muodon avulla voidaan tehdä jako mm. julkisen ja yksityisen sektorin välillä. Tieto juridisesta muodosta perustuu laskentalomakkeella saatuihin tietoihin henkilön työpaikasta.

**Taulukko 12: Työpaikan juridinen muoto**

Työpaikan juridinen muoto	Vanha tieto	Uusi tieto	Sama luok. %	Nettovirhe %
Julkinen sektori	463	464	96.8	-0.2
Valtio	155	157	95.5	-1.3
Kunta	243	242	94.6	0.4
Valtioenemmistöinen osakeyhtiö	65	65	95.4	0.0
Yksityinen sektori	961	963	97.4	-0.2
Muu osakeyhtiö, paitsi asunto-osakeyhtiö	491	496	97.1	-1.0
Muu yksityinen sektori	470	467	94.8	0.6
Muu, tuntematon, työpaikka ulkomailta	31	28	57.6	10.7
Ei työllinen	6	7	(0)	..
<b>Yhteensä</b>	<b>1461</b>	<b>1461</b>	<b>94.5</b>	<b>-</b>

94.5 % kaikista tapauksista sai molemmilla kerroilla saman juridisen muodon. Poikkeavia tapauksia oli yhteensä 81. Mikäli vain toimipaikat otetaan mukaan tarkasteluun, saman koodin saaneiden osuus nousee hiukan (95.9 %).

Mikäli työpaikat jaetaan kolmeen luokkaan 'Julkinen sektori', 'Yksityinen sektori' ja 'Muu' samalla tavalla luokitelujen tapausten osuus kasvaa 96.0 %:iin.

Noin puolet poikkeavista tapauksista johtui siitä, että uusintatutkimuksen juridinen muoto oli virheellinen. Näin varsinaisia virhetapauksia alkuperäisaineistossa oli noin 40 (3 % työllisistä). Taulukon 12 lukuja samalla tavalla luokiteltujen

osuuksista voidaan näin pitää minimiarvoina -  
vertailu korjattuun arvoon antaisi korkeampia  
prosenttiosuuksia. Virheiden syitä ei ole tarkemmin  
selvitetty.

## 6. VERATILU VUODEN 1980 LASKENNAN KÄSITTELYVIRHESELVITYKSEN TULOKSIIN

Vuoden 1980 käsittelyvirheselvitys koski yhteensä 14 tietoa, joiden joukossa oli myös rakennuksia koskevia muuttujia. Seuraavat tiedot olivat mukana sekä vuoden 1980 että vuoden 1985 selvityksissä:

- ammatti
- elinkeino (toimiala)
- liikkuvuustyyppi
- työpaikan sijaintikunta
- työpaikan juridinen muoto

Seuraavassa keskitytäänkin näihin muuttujiin.

Vuoden 1980 laskennan lomakekäsittely poikkesi monelta kohdin vuoden 1985 ratkaisusta. Monet tiedot esim. kooditettiin suoraan lomakkeelle kuntiin perustetuissa paikallistoimistoissa. Ns. henkilölomakkeelle kooditettiin ammatti, työpaikkaa koskevat tiedot (sijaintikunta, toimiala, juridinen muoto ja liikkuvuustyyppi) sekä oppilaitostunnus. Carita Putkonen (1983, 97) kuvaa käsittelyvirheselvityksen kulkua seuraavasti:

Otokseen kuuluneet --- lomakkeet kopioitiin --- siten, että paikallistoimistossa tehdyt koodimerkinnät poistettiin. Tämän jälkeen kopioidut lomakkeet merkittiin keskusyksikössä. Koodatuille lomakkeille suoritettiin VAL 80:n oikeellisuus- ja ristiriidattomuustarkistukset. Selvityksen korjattuja tietoja verrattiin koneellisesti VAL 80:n vastaaviin tietoihin. Poikkeavat tiedot tulostettiin virhelistalle, jolle tulostettu käsittelyvirheselvityksen koodi tarkistettiin. --- Virhelistalle korjatut tiedot tallennettiin ja tallennuksen jälkeen tarkistettiin vielä mahdolliset tallennusvirheet. ---

Käsittelyvirheselvityksen lopullisia tietoja verrattiin yksikkökohtaisesti VAL 80:n lopullisiin tietoihin ja laskettiin ns. bruttovirheet. --.

Käsittelyvirheselvityksessä oli mukana yhteensä 3 686 henkilölomaketta.

Itse tallennuksen laatua tutkittiin vertailemalla kahteen kertaan tallennettuja tietoja keskenään. Saatujen tulosten mukaan tallennusvirheitä oli alle prosentissa tutkituista tapauksista. Eri lomaketyyppien tallennusvirheiden määrät eivät poikenneet toisistaan. Myös tiedoittain tallennusvirheiden määrät vaihtelivat hyvin vähän. Ammattikoodin kohdalla virheellisten tallennustietojen osuus oli 0.5 %, toimialakoodin kohdalla 0.6 % (koko otoksesta). Yleisin tallennusvirhe oli tiedon jättäminen kokonaan tallentamatta.

Näitä lukuja ei voida suoraan verrata vuoden 1985 ammatti- ja toimialatietoihin. Silloinhan tallennettiin selväkielinen ammatti ja/tai toimialatieto. Sen sijaan tulosta voitaneen verrata 1985 laskentalomakkeen numeerisiin tallennustietoihin. Näissä virheellisten tallennustietojen osuus vaihteli 0.3 ja 2.7 % välillä. Keskimäärin poikkeavia numerokodeja oli 0.6 %. Tallennuksen laatu näyttää siis olleen suunnilleen yhtä hyvä vuonna 1980 ja vuonna 1985.

Mitä tulee lopullisten muuttujien 'bruttovirheisiin' (100 miinus samalla tavalla luokiteltujen osuus) vertailun tulokset näkyvät taulukosta 13.

**Taulukko 13: Muuttujien bruttovirheet vuonna 1980 ja 1985**

Tieto	1980 %	1985 %
Pääasiallinen toiminta	..	0.8
Kokopäivätyökuukaudet	..	0.4
Osa-aikatyökuukaudet (väh. 20 h/v)	..	0.4
Työttömyyskuukaudet	..	0.3
Ammatissatoimivuus	..	0.3
Ammattiasema	..	1.7
Ammatti (4-numerotaso)	14.1	11.6
Elinkeino (5-numerotaso)	11.9	14.1
Työpaikan sijaintikunta	3.4	5.6
Liikkuvuustyyppi	7.1	8.4
Työpaikan juridinen muoto	6.5	5.5

Vuoden 1980 tiedot: Putkonen (1983, 115)

Taulukosta 13 nähdään, että bruttovirheiden osuus vuonna 1985 joskus on pienempi ja joskus taas suurempi kuin vuonna 1980. Vaikka tietojen käsittely vuonna 1985 monilta osin poikkesi vuoden 1980 ratkaisuksista, selviä eroja ei ole havaittavissa. Työpaikan sijaintikunta -tieto näyttää olevan hie- man luotettavampi vuonna 1980 - silloinhan jokai-

sessä kunnassa oli oma laskentatoimisto. Toisaalta taas näyttää siltä, että ammattien esitäyttö ja automaattinen kooditus on ollut omiaan parantamaan ammattitiedon laatua.

Vertailuja tehdessä on kuitenkin otettava huomioon, että vuoden 1980 käsittelyvirheselvityksessä uusintatutkimuksen tiedot korjattiin, mikäli ne havaittiin virheellisiksi. Vuoden 1985 aineistossa korjattiin vain tallennusvirheet. Vuoden 1985 lukuja voidaan siten pitää bruttovirheen maksimiarvoina - todellinen bruttovirhe on pienempi. Lisäksi luokiteltiin vuonna 1980 maa- ja metsätaloustyössä olleita (toimialaluokka 1) tarkemmin kuin vuonna 1985, jolloin tarkin luokitustaso oli 2-numeroinen. Juridinen muoto jaoteltiin myös tarkemmin vuonna 1980 (15 luokkaa) kuin vuonna 1985 (6 luokkaa).

Yhteenvetona voitaneen kuitenkin sanoa, että käsittelyvirheiden osuus vuoden 1985 laskennassa ei ylittänyt vuoden 1980 tasoa - ainakaan kun tarkastelun kohteena ovat koko maan tiedot. Tästä syystä vuoden 1985 selvästi halvemmalle ratkaisulle voitaneen antaa hyväksytty arvosana.

## 7. KÄSITTELYVIRHEIDEN SEURAUKSET

Mikäli käsittelyvirheet ovat luonteeltaan satunnaisia ne kumoavat toisiaan siten, että muuttujien reunajakaumat eivät muutu. Satunnaisvirheiden seurauksena kahden muuttujan välinen riippuvuussuhde saattaa kuitenkin hämärtyä ja jäädä peittoon. Erityisesti liikkuvuutta tutkittaessa satunnaisvirheiden merkitys korostuu.

Kun pitkittäisaineiston avulla tutkitaan, kuinka moni henkilö tietyssä ajanjaksona on siirtynyt ryhmästä toiseen, esim. vaihtanut ammattia, on tärkeätä tietää kuinka suuri osa havaituista eroista johtuu luokitusvirheistä ja kuinka suuri osa taas kuvaa todellista liikkuvuutta. Mikäli esim. vuoden 1985 laskentatietojen avulla halutaan tutkia ammatillista liikkuvuutta 1-numerotasolla, on syytä muistaa, että noin 7 % perusjoukosta "vaihtoi ammattia" jo silloin, kun samat lomaketiedot kooditettiin kaksi kertaa toisistaan riippumatta.

Taulukossa 14 onkin laskettu, kuinka moni henkilö luokiteltiin samaan ammattiryhmään vuonna 1980 ja vuonna 1985. Vuoden 1985 tietoina - ja prosenttien peruslukuna - on käytetty sekä varsinaisen laskennan tuloksia että uusintatutkimuksen ammatitietoja.

**Taulukko 14: Ammattijakauma 1980 - 1985 sekä ammatillinen liikkuvuus**

Ammattiryhmä	1980	(a) 1985	(b) 1985	(c) sama	(d) sama	c/a %	d/b %
<b>Yhteensä</b>	2188	2188	2188	1465	1475	67.0	67.4
0 Tekninen työ	257	311	307	197	202	63.3	65.8
1 Hallinnollinen työ	190	211	219	132	137	62.6	62.6
2 Kaupallinen työ	100	105	100	54	53	51.4	53.0
3 Maa- ja metsätaloustyö, kalastusala	148	143	148	97	97	67.8	65.5
4 Kaivos-, syväkairaus- ja rikastustyö	5	4	5	3	2	..	..
5 Kuljetus- ja liikennet.	86	93	95	55	57	59.1	60.0
6/7 Teollinen työ, koneen- hoito ym.	406	419	411	297	296	70.9	72.0
8 Palvelutyö	155	162	158	98	98	60.4	62.0
9 Muualla luokittelematon työ ja työvoima	5	7	11	1	1	..	..
Ei työllinen	820	733	734	522	523	71.2	71.2
Ei mukana vuonna 1980	16						
2-numerotaso	2188	2188		1324	1338	60.5	61.2
3-numerotaso	2188	2188		1225	1243	56.0	56.8
4-numerotaso	2188	2188		1217	1232	55.6	56.3

a = ammattiryhmän koko 1985 alkuperäisen aineiston mukaan

b = ammattiryhmän koko 1985 uusintatutkimuksen mukaan

c = sama ammattiryhmä 1980 ja vuoden 1985 alkuperäisessä aineistossa

d = sama ammattiryhmä 1980 ja vuoden 1985 uusintatutkimuksen aineistossa

Taulukosta 14 voidaan nähdä, että luvut eivät juuri poikkea toisistaan. Joitakin pieniä eroja on kuitenkin havaittavissa. Esimerkiksi niistä henkilöistä, jotka varsinaisen laskenta-aineiston mukaan toimivat teknisissä ammateissa vuonna 1985, oli 63.3 % luokiteltu samaan ammattiryhmään vuonna 1980. Mikäli sama vertailu tehdään vuoden 1985 uusintatutkimuksen aineistolla vastaava prosenttiosuus on 65.8 %. Mikäli myös vuoden 1980 laskennan ammateista olisi käytettävissä kaksi eri versiota, liikkuvuustiedot saattaisivat poiketa toisistaan vielä selvemmin.

Liitetaulukoissa 8a ja 8b on esitetty yksityiskohdalliset matriisit ammatillisesta liikkuvuudesta.

## 8. LOPUKSI

Väestölaskennoissa kohdehenkilön antamat tiedot ovat lähinnä "aputietoja", jotka yhdessä muiden käytettävissä olevien tietojen kanssa muodostavat pohjan mahdollisimman "oikeata" koodiarvoa etsittäessä. Näin laskennat eroavat esim. vaalitulastoista, joita tehtäessä ei tarvitse pohtia, mitä puoluetta henkilön "olisi pitänyt" äänestää.

Edellä mainitusta syystä laskentatilastojen luotettavuutta ei voida arvioida siten, että henkilön saamaa koodiarvoa yksinkertaisesti verrataan hänen itsensä antamaan tietoon. Vuoden 1985 laskennassa ei ollut yhtään sellaista kysymystä, jossa henkilön itse antamaa tietoa määritelmällisesti voitaisiin pitää oikeana. Esimerkkejä tällaisesta kysymyksestä löytyy kyllä aikaisemmista laskennoista:

Pääkieli. Epäselvissä tapauksissa ilmoittakaa mihin kieliryhmään lähinnä katsotte kuuluvanne. --

(Vuoden 1970 laskennan kielikysymys)

Käsittelyvirheselvityksen tehtävä ei siten ole vain sen tutkiminen, kuinka luotettavasti lomaketiedot tallennettiin (tällöinhän periaate oli, että tallennettu versio olisi 1/1-kopio lomakkeesta) tai paljonko sattuma pääsi vaikuttamaan tietojen koodituksessa. Käsittely(virhe)selvityksen pitäisi myös antaa tietoa siitä, kuinka usein kohdehenkilö, joka oli palauttanut harhaanjohtavan ja puutteellisen lomakkeen, onnistuttiin sijoittamaan "oikeaan" lokeroon ja - vastaavasti - kuinka usein "oikea" lomaketieto syystä tai toisesta tuli hylätyksi. Onhan lomakeaineiston muokkaaminen osa käsittelyä.

Viimeksimainittuun ongelmaan ei kuitenkaan kajota tässä tutkimuksessa. Lomaketietojen oikeellisuutta ja oikeellisuuteen vaikuttavia tekijöitä tutkitaan tarkemmin erillisessä luotettavuustutkimuksessa, jossa tietoja verrataan Tilastokeskuksen työvoimatutkimuksen yhteydessä kerättyyn aineistoon, jota voidaan pitää erittäin luotettavana.

Eräs tarkastelunäkökulma laskentatietojen laatua ja luotettavuutta tutkittaessa on verrata lomakkeella kerättyjä tietoja vastaaviin muista lähteistä saatuihin tietoihin. Näin saadaan käsitys siitä, kuinka paljon tilastotuotannossa käytetyt tiedot ovat riippuvaisia tietyistä käsittely- ja mittausmistavoista.

Tilastokeskuksessa on tutkittu rekisteripohjaisen laskennan mahdollisuuksia ja ongelmia mm. vertailemalla vuoden 1985 väestölaskennan tuloksia ns. rinnakkaistilastoon, jossa kaikki yli 15-vuotiaat väestölaskennan perusjoukkoon kuuluvat luokiteltiin eri rekisteritietoja hyväksikäyttäen pääasiallisen toiminnan, ammattiaseman ja työpaikan juridisen muodon mukaan. Taulukko 15 antaa kuvan vertailun tuloksista.

**Taulukko 15: Pääasiallinen toiminta, ammattiasema ja työpaikan juridinen muoto vuoden 1985 väestölaskennan sekä rekisteripäätelyn mukaan**

Pääasiallinen toiminta, ammattiasema, jur. muoto	VL 85 -tieto	Rekisteripäätely	Sama luok. %	Nettovirhe %
Työvoima	2415 991	2452 843	94.7	1.5
Työllinen työvoima	2276 887	2318 757	94.4	1.8
Palkansaaajat	1956 136	1997 215	93.9	2.1
Valtio	230 135	235 778	87.2	2.5
Kunta	400 416	438 358	87.3	9.5
Yksityinen sektori	1325 585	1323 079	91.7	-0.2
Yrittäjät ja avustavat perheenjäsenet	320 685	321 542	79.3	0.3
Työnantajat	50 358	30 044	45.0	-40.3
Yksinäisyrittäjät	178 354	231 271	64.7	29.7
Avust. perh.jäs.	91 973	60 227	57.4	-34.5
Työtön työvoima	139 104	134 086	70.9	-3.6
Työvoimaan kuulumattomat	1543 258	1506 406	91.5	-2.4
Opiskelijat, koululaiset	370 260	297 990	82.8	-19.5
Kotitaloustyötä tekevät	99 825	..	..	..
Eläkeläiset	938 202	957 128	95.6	2.0
Varusmiehet	28 688	..	..	..
Muut työvoiman ulkopuolella olevat	38 969	..	..	..
Tuntematon	67 314	..	..	..
Yhteensä (väestön pääasiallinen toiminta)	3959 249	3959 249	90.3	-
Yhteensä (työllisten ammattiasema)	2467 170	2467 170	83.1	-
Yhteensä (palkansaaajien työpaikan jur. muoto)	2098 139	2098 139	85.1	-

Lähde: Korpi 1988 (toistaiseksi julkaisematon aineisto)

Taulukosta 15 nähdään, että samalla tavalla luokiteltujen osuus vaihtelee 45.0 %:in (työnantajat) ja 94.7 %:in (työvoima) välillä. Keskimäärin noin joka kymmenes henkilö sai rinnakkaistilastoissa eri koodiarvon kuin väestölaskennassa. Poikkeavat tapaukset johtuvat joko virheistä tai siitä, ettei rinnakkaistilastoissa onnistuttu löytämään päätte-

lysäntöjä, jotka vastaisivat vuoden 1985 väestölaskennan mittausoperaatioita.

Taulukosta 15 nähdään myös selvästi, että varsinkin yrittäjien jakaminen väestölaskennan 1985 mukaisiin osaryhmiin pelkästään rekisteritietoja hyväksikäyttäen on vaikeata. Yrittäjien ja avustavien perheenjäsenten välistä erottelua ei ole onnistuttu toistamaan käytettävissä olevien rekisteritietojen ja päättelysääntöjen avulla. Sellaiset luokitukset, joita ei juuri käytetä muualla kuin tilastotuotannossa näyttävät olevan ongelmallisempia kuin luokitukset, joita käytetään muissakin yhteyksissä ja jotka siten selvemmin muodostavat "sosiaalisia kategorioita".

## KIRJALLISUUSLUETTELO

Luettelo sisältää kaikki tärkeimmät kirjoitukset, jotka liittyvät vuoden 1985 väestölaskentaan tai joissa käsitellään väestölaskentojen luotettavuuteen liittyviä kysymyksiä.

Kauko Erkko: Väestölaskennan ja työvoimatiedustelun tulosten vertailemistä. Tilastokeskus 22.01.1973.

Risto Heinonen ja Aarno Laihonen: Some new solutions and methods for census data production: Finnish experiences from the 1985 Census. Paper prepared for the ECE/CES Seminar on Computer-related Aspects of Population and Housing Censuses, Belgrad, October 1987.

Salme Kiiski: Työvoimatutkimuksen luotettavuusselvitys väestölaskennan 1985 luotettavuustutkimuksen yhteydessä. Tilastokeskus, Raportti 1988:4.

Helena Korpi: Pääasiallisen toiminnan, toiminnan laadun ja ammattiaseman määrittely rekisterien avulla: kunnittainen koetilasto. Tilastokeskus, Raportti 1986:17.

Helena Korpi: Vuoden 1985 väestölaskennan rekisteripohjainen rinnakkaistilasto (käsikirjoitus).

L-M. Laaksonen: Tutkintorekisterin tietojen tarkistus HORGAn kokoomatutkimuksen 1/1977 ja 2/1977 yhteydessä. Tilastokeskus, koulutustilastotoimisto 26.8.1977.

Aarno Laihonen ja Pekka Myrskylä: Use of registers and administrative records in population censuses in Finland. Paper prepared for the European Population Conference, Jyväskylä, June 11-16, 1987.

T. Lehtonen: Rakennuskoordinaattien tarkkuuden tutkimus. Maanmittaushallitus, 12.7.1973.

Mauri Levomäki: Raportti VAL 80:n tiedotuksesta. Tilastokeskus, Raportti 1982:11.

Eyvor Lindgren ja Jari Tarkoma: Raportti AET:n koetiedustelusta 1.12.1984. Tilastokeskus 19.3.1985.

Pasi Markelin: Pääasiallinen toiminta-käsite (Type of Activity) väestölaskennoissa. Ammatissa toimiva väestö/työvoima. Tilastokeskus, Muistio 42 (1977).

Pasi Markelin: Peittävyys selvitys, Asunto- ja elinkeinotutkimus 1975. Tilastotiedotuksia VÄ 1978:20.

Pasi Markelin: Laatuongelmat rekistereitä käytettäessä, julkaisussa Tilastojen laadun tutkiminen, Tilastokeskus, Tutkimuksia nro 104 (1984).

Pekka Myrskylä: Vuoden 1980 väestö- ja asuntolaskenta. Raportti Toijalan ja Luumäen koelaskennasta. Tilastokeskus, VAL 80, 6.4.1979.

Erkki Pentinmäki: Väestö- ja asuntolaskentaa koskeva mielipidetutkimus Kauniaisissa ja Särkisalossa tammikuussa 1969. Tilastollinen päätoimisto 14.7.1969.

Erkki Pentinmäki: Väestölaskentatietojen luotettavuus. Demografian jatkokoulutusseminaari 1977, Suomen väestötieteen yhdistyksen julkaisuja N:o 3 (1978).

Carita Putkonen: VAL 80:n luotettavuustutkimus 9.5.1980. Tilastokeskus, Väestö- ja asuntolaskenta 1980. (Sisältää tutkimussuunnitelman lisäksi liitteen: Väestölaskentojen yhteydessä suoritettut luotettavuustutkimukset).

Carita Putkonen: Luotettavuustutkimus, Väestö- ja asuntolaskenta 1980, Osa I, Selvitys mittaus- ja käsittelyvirheistä. Tilastokeskus, Tutkimuksia nro 99 (1983).

Carita Putkonen: Confidence Study. Population and Housing Census 1980. Tilastokeskus, Tutkimuksia (Studies) 99 (1984). (Edellisen julk. engl. käännös).

Carita Putkonen: Luotettavuustutkimus, Väestö- ja asuntolaskenta 1980, Osa II, Peittävyys selvitys. Tilastokeskus, Tutkimuksia nro 100 (1984).

Carita Putkonen: Väestölaskennan laatututkimus; esimerkki laadun mittaamisesta, julkaisussa Tilastojen laadun tutkiminen.

Christian Starck: Väestölaskenta 1985: Tiedotusraportti. Tilastokeskus, Muistio 1987:106.

Christian Starck: Väestölaskenta 1985: Lomakekeruun loppuraportti. Tilastokeskus, Muistio 1988:116.

VL 90-projekti. Suunnitelma rekisteripohjaisesta väestölaskentajärjestelmästä. Tilastokeskus, Raportti 1987:4.

Lauri Yli-Pekka ja Risto Heinonen: Vuoden 1985 väestölaskennan loppuraportti. Tilastokeskus, Raportti 1988:7.

Yrittäjäkäsité tilastoissa. Tilastokeskus, Raportti 1986:16.

Vuoden 1985 väestölaskennasta on tehty viisi julkaisua Suomen virallinen tilasto -sarjassa:

- I Väestön taloudellinen toiminta
- II Asuntokunnat ja perheet
- III Asuinolot
- IV Rakennukset ja asunnot
- V Kunnittaiset tiedot

**Liitetaulukko 1: Pääasiallinen toiminta varsinaisen aineiston ja uusintatutkimuksen aineiston mukaan**

Varsinainen aineisto	Uusintatutkimus								Yht.
	11	12	22	23	24/29	25	26	99	
11 Työllinen työvoima	1448	1	0	3	2	0	1	0	1455
12 Työtön työvoima	0	85	0	0	0	0	1	0	86
22 Opiskelijat, koululaiset	1	0	212	0	0	0	0	0	213
23 Kotitaloustyötä tekevät	3	0	0	58	0	0	0	0	61
24/29 Eläkeläiset	2	0	0	1	326	0	0	0	329
25 Varusmiehet	0	0	0	0	0	20	0	0	20
26 Muut työvoiman ulkopuolella olevat	0	0	0	1	0	0	21	0	22
99 Tunteaton	0	0	0	0	0	0	2	0	2
<b>Yhteensä</b>	<b>1454</b>	<b>86</b>	<b>212</b>	<b>63</b>	<b>328</b>	<b>20</b>	<b>25</b>	<b>0</b>	<b>2188</b>

**Liitetaulukko 2: Ammattiasema varsinaisen aineiston ja uusintatutkimuksen aineiston mukaan**

Varsinainen aineisto	Uusintatutkimus					Yht.
	1	2	3	4	Ei t.	
1 Palkansaaja	1235	0	4	1	2	1242
2 Yksinäisyrittäjä	2	113	1	1	1	118
3 Yrittäjä, jolla palkattua työvoimaa	0	0	42	0	1	43
4 Yrittäjän avustava perheenjäsen	0	1	2	46	3	52
Ei työllinen	1	2	0	3	727	733
<b>Yhteensä</b>	<b>1238</b>	<b>116</b>	<b>49</b>	<b>51</b>	<b>734</b>	<b>2188</b>

**Liitetaulukko 3: Ammatissa toimivuus varsinaisen aineiston ja uusintatutkimuksen aineiston mukaan**

Varsinainen aineisto	Uusintatutkimus			
	1	2	3	Yht.
1 Ammatissa toimiva	1419	2	1	1422
2 Ammatissa toimimaton	3	756	0	759
3 Tunteaton	1	0	6	7
<b>Yhteensä</b>	<b>1423</b>	<b>758</b>	<b>7</b>	<b>2188</b>

**Liitetaulukko 4: Ammatti (1-nro) varsinaisen aineiston ja uusintatutkimuksen aineiston mukaan**

Vars. aineisto	Uusintatutkimus										Yht.
	0	1	2	3	4	5	6/7	8	9	Ei t	
0 Tekninen työ	291	9	1	0	0	1	5	3	1	0	311
1 Hallinnollinen työ	7	198	2	0	0	0	2	1	0	1	211
2 Kaupallinen työ	1	6	94	0	0	0	0	2	1	1	105
3 Maa- ja metsätaloustyö, kalastusala	0	0	0	140	0	0	0	0	1	2	143
4 Kaivos-, syväkairaus- ja rikastustyö	0	0	0	0	3	0	0	0	0	1	4
5 Kuljetus- ja liikennet.	0	1	0	0	0	91	1	0	0	0	93
6/7 Teollinen työ, koneen- hoito ym.	7	2	3	2	2	2	401	0	0	0	419
8 Palvelutyö	0	3	0	1	0	0	2	152	3	1	162
9 Muualla luokittelematon työ ja työvoima	1	0	0	0	0	1	0	0	4	1	7
Ei työllinen	0	0	0	5	0	0	0	0	1	727	733
<b>Yhteensä</b>	<b>307</b>	<b>219</b>	<b>100</b>	<b>148</b>	<b>5</b>	<b>95</b>	<b>411</b>	<b>158</b>	<b>11</b>	<b>734</b>	<b>2188</b>

**Liitetaulukko 5: Elinkeino (1-nro) varsinaisen aineiston ja uusintatutkimuksen aineiston mukaan**

Vars. aineisto	Uusintatutkimus										Yht.
	1	2	3	4	5	6	7	8	9	Ei t	
1 Maatalous	140	0	0	0	0	0	0	0	2	2	144
2 Kaivostoiminta	0	7	1	0	0	0	0	0	0	0	8
3 Teollisuus	1	1	361	0	3	1	2	1	0	0	370
4 Sähkö- yms. huolto	0	0	1	25	1	0	0	0	0	0	27
5 Rakennustoiminta	0	0	1	0	94	0	0	1	6	2	104
6 Kauppa	2	0	1	0	0	178	1	2	6	1	191
7 Kuljetus	0	0	0	2	0	0	101	1	1	0	105
8 Rahoitus	0	0	1	0	0	0	0	96	3	1	101
9 Yhteisk. palvelut	1	0	2	0	1	4	2	0	394	0	404
Ei työllinen	5	0	1	0	0	0	0	0	0	728	734
<b>Yhteensä</b>	<b>149</b>	<b>8</b>	<b>369</b>	<b>27</b>	<b>99</b>	<b>183</b>	<b>106</b>	<b>101</b>	<b>412</b>	<b>734</b>	<b>2188</b>

**Liitetaulukko 6: Liikkuvuustyyppi varsinaisen aineiston ja uusintatutkimuksen aineiston mukaan**

Varsinainen aineisto	Uusintatutkimus									Yht.
	0	1	2	3	4	5	7	9	Ei t	
0 Työpaikka asunnon yhteydessä	130	3	0	0	2	0	1	1	4	141
1 Yhdessä kunnassa liikkuva	4	55	0	5	15	0	8	4	1	92
2 Kunnassa tuntematon	2	0	1	1	0	0	8	4	0	16
3 Useissa kunnissa liikkuva	0	1	0	11	0	0	2	1	0	15
4 Kiinteä työpaikka, mutta ei toimipaikka	0	2	0	1	28	0	1	1	0	33
5 Ulkomailla	0	0	0	0	0	3	0	0	0	3
7 Toimipaikka	19	14	2	1	2	0	1107	4	1	1150
9 Tuntematon	1	1	0	0	0	0	1	1	1	5
Ei työllinen	5	0	0	0	0	0	1	0	727	733
<b>Yhteensä</b>	<b>161</b>	<b>76</b>	<b>3</b>	<b>19</b>	<b>47</b>	<b>3</b>	<b>1129</b>	<b>16</b>	<b>734</b>	<b>2188</b>

**Liitetaulukko 7: Juridinen muoto varsinaisen aineiston ja uusintatutkimuksen aineiston mukaan**

Varsinainen aineisto	Uusintatutkimus							Yht.
	01	02	03	04	05	09	Ei t.	
01 Valtio	149	1	1	0	2	1	1	155
02 Kunta	5	229	0	0	4	4	1	243
03 Valtioenemmistöinen osakeyhtiö	1	0	62	2	0	0	0	65
04 Muu osakeyhtiö, paitsi asunto-osakeyhtiö	0	1	2	479	7	2	0	491
05 Muu yksityinen sektori	2	8	0	7	444	5	4	470
09 Muu, tuntematon, työpaikka ulkomailla	0	1	0	8	5	17	0	31
Ei työllinen	0	1	0	0	5	0	727	733
<b>Yhteensä</b>	<b>157</b>	<b>241</b>	<b>65</b>	<b>496</b>	<b>467</b>	<b>28</b>	<b>734</b>	<b>2188</b>

Liitetaulukko 8a: Ammatillinen liikkuvuus 1980 - 1985 (varsinainen laskenta-aineisto)

Ammattiryhmä 1985  
(varsinainen aineisto)

	Ammattiryhmä vuonna 1980																			Yhteensä
	1)	11	12	13	14	15	16/7	18	19	1	2	3	4	5	6	7	8	9	10	
Ei työllinen	531	24	20	19	36	0	11	60	31	1										733
työllinen	72.44	3.27	2.73	2.59	4.91	0.00	1.50	8.19	4.23	0.14										
0	66	197	15	4	4	0	3	15	7	0										311
	21.22	63.34	4.82	1.29	1.29	0.00	0.96	4.82	2.25	0.00										
1	34	20	132	6	1	0	6	6	5	1										211
	16.11	9.48	62.56	2.84	0.47	0.00	2.84	2.84	2.37	0.47										
2	24	5	9	54	1	0	2	6	4	0										105
	22.86	4.76	8.57	51.43	0.95	0.00	1.90	5.71	3.81	0.00										
3	33	1	2	1	97	0	3	6	0	0										143
	23.08	0.70	1.40	0.70	67.83	0.00	2.10	4.20	0.00	0.00										
4	1	0	0	0	0	3	0	0	0	0										4
	25.00	0.00	0.00	0.00	0.00	75.00	0.00	0.00	0.00	0.00										
5	22	1	2	4	2	0	55	6	1	0										93
	23.66	1.08	2.15	4.30	2.15	0.00	59.14	6.45	1.08	0.00										
6/7	81	3	6	8	7	2	4	297	9	2										419
	19.33	0.72	1.43	1.91	1.67	0.48	0.95	70.88	2.15	0.48										
8	41	5	3	4	0	0	1	10	98	0										162
	25.31	3.09	1.85	2.47	0.00	0.00	0.62	6.17	60.49	0.00										
9	3	1	1	0	0	0	1	0	0	1										7
	42.86	14.29	14.29	0.00	0.00	0.00	14.29	0.00	0.00	14.29										
Yhteensä	836	257	190	100	148	5	86	406	155	5										2188

1) Sisältää 16 henkilöä, jotka eivät olleet mukana vuoden 1980 perusjoukossa.

Liitetaulukko 8b: Ammatillinen liikkuvuus 1980 - 1985 (uusintatutkimuksen aineisto)

Ammattiryhmä 1985 (uusintatutkimuksen aineisto)		Ammattiryhmä vuonna 1980																		
		1)																		
Ei työl.		10	11	12	13	14	15	16/7	18	19										Yhteensä
Ei työllinen	532	24	21	19	35	1	11	60	30	1										734
	72.48	3.27	2.86	2.59	4.77	0.14	1.50	8.17	4.09	0.14										
0	62	202	12	3	5	0	2	15	6	0										307
	20.20	65.80	3.91	0.98	1.63	0.00	0.65	4.89	1.95	0.00										
1	40	17	137	8	1	0	6	5	4	1										219
	18.26	7.76	62.56	3.65	0.46	0.00	2.74	2.28	1.83	0.46										
2	22	4	9	53	1	0	2	6	3	0										100
	22.00	4.00	9.00	53.00	1.00	0.00	2.00	6.00	3.00	0.00										
3	34	1	2	2	97	0	3	8	1	0										148
	22.97	0.68	1.35	1.35	65.54	0.00	2.03	5.41	0.68	0.00										
4	1	0	1	0	0	2	0	1	0	0										5
	20.00	0.00	20.00	0.00	0.00	40.00	0.00	20.00	0.00	0.00										
5	20	1	2	4	2	0	57	6	3	0										95
	21.05	1.05	2.11	4.21	2.11	0.00	60.00	6.32	3.16	0.00										
6/7	79	4	4	7	6	2	4	296	7	2										411
	19.22	0.97	0.97	1.70	1.46	0.49	0.97	72.02	1.70	0.49										
8	41	4	1	4	0	0	1	9	98	0										158
	25.95	2.53	0.63	2.53	0.00	0.00	0.63	5.70	62.03	0.00										
9	5	0	1	0	1	0	0	0	3	1										11
	45.45	0.00	9.09	0.00	9.09	0.00	0.00	0.00	27.27	9.09										
Yhteensä	836	257	190	100	148	5	86	406	155	5										2188

1) Sisältää 16 henkilöä, jotka eivät olleet mukana vuoden 1980 perusjoukossa.



Tilastokeskus

LAKISÄÄTEINEN VÄESTÖLASKENTA 1985

Luottamuksellinen

Tilastokeskus suorittaa marraskuun 17. päivänä lakiin perustuvan väestölaskennan.

Kaikkien 18.11.1910-17.11.1970 syntyneiden on täytettävä tämä lomake (laki ammatti- ja elinkeinotutkimuksesta 377/85, asetus 440/85).

Tilastokeskus on valtion virasto, jonka tehtävänä on laatia yhteiskuntaa kuvaavia tilastoja. Väestölaskennoilla kuvataan muun muassa väestön työssäkäyntiä sekä elinkeino- ja ammattirakennetta.

Lomakkeella kerättäviä tietoja käytetään yksinomaan tilastollisiin tarkoituksiin. Lain mukaan Tilastokeskus pitää yksityistä henkilöä koskevat tiedot salassa.

Palauttakaa lomake oheisessa palautuskuoressa ilman postimaksua viimeistään 22.11.1985.

Nimi

Osoite

① Henkilötunnus  
(sosiaaliturvatunnus)

Syntymäpäivä kuukausi vuosi-tunnusosa

② Toimintanne viikolla 11.11.-17.11. 1985

Merkitkää rasti (x) Teihin sopivan vaihtoehdon kohdalle. Jos olitte ansiotyössä yhtenäkin päivänä tai tilapäisesti poissa työstä loman, sairauden tai muun sellaisen vuoksi, valitkaa jokin kohdista 1-4.

- 1  Palkansaaja
- 2  Yksinäisyrittäjä tai ammatinharjoittaja  
(ei palkattua työvoimaa)
- 3  Työnantajayrittäjä tai ammatinharjoittaja  
(palkattua työvoimaa)
- 4  Yrittäjän avustava perheenjäsen (katso ohjetta)
- 5  Työtön tai pakkolomalla
- 6  Eläkeläinen
- 7  Opiskelija tai koululainen
- 8  Varusmies tai siviilipalvelusmies
- 9  Omaa kotitaloutta hoitava
- 0  Muu, mikä? \_\_\_\_\_

#### OHJEET

**Ansiotyöllä** tarkoitetaan kaikkea tuloa tuottavaa työtä. Myös ilman varsinaista palkkaa tehty työ perheenjäsenen maatilalla tai yrityksessä katsotaan ansiotyöksi. Jos teitte päätyön ohessa sivutyötä, antakaa tiedot päätyöpaikan mukaan.

**Palkansaajia** ovat kaikki, jotka tekevät toisen palveluksessa työtä palkkaa tai palkkiota vastaan. Palkansaajiksi katsotaan myös apurahan saaja sekä henkilö, joka työskentelee päätoimisesti osakeyhtiössä, jonka osakas hän on.

**Yrittäjä tai ammatinharjoittaja** on henkilö, joka hoitaa omaa yritystään, viljelee omistamaansa tai vuokraamaansa maatilaa taikka harjoittaa itsehäistä (esim. ompelijan tai kirjailijan) ammattia.

**Yrittäjän avustava perheenjäsen** on henkilö, joka ilman varsinaista palkkaa työskentelee samassa taloudessa asuvan yrittäjän yrityksessä.

**Työttömällä** tarkoitetaan henkilöä, joka on työtä vailla, etsii työtä ja on työhön käytettävissä tai on lomautettu palkatta.

**Eläkeläisiä** ovat vanhuus-, työ-, työkyvyttömyys-, työttömyys-, luopumis-, sukupolvenvaihdos- tai rintamasotilaseläkkeellä olevat ja syyttingillä tai vanhuudentuella elävät. Sen sijaan perhe-eläkkeen varassa elävät rasti-tavat kohdan "Muu".

**Opiskelulla tai koulunkäynnillä** tarkoitetaan päätoimista opiskelua oppilaitoksessa sekä opiskeluun liittyvää palkatonta työharjoittelua.

**Muiksi toiminnaksi** luetaan omaisuustuloilla, säästöillä tai perhe-eläkkeellä eläminen ja muu edellisiin ryhmiin kuulumaton toiminta, josta ei makseta palkkaa.

**NE JOTKA MERKITSIVÄT EDELLISESSÄ KYSYMYKSESSÄ RASTIN KOHTAAN 1, 2, 3 TAI 4, VASTAAVAT JATKOKYSYMYKSIIN. MUUT SIIRTYVÄT SUORAAN KYSYMYKSEEN 8.**

(A) ALLA OVAT VUODEN 1980 VÄESTÖLASKENNAN TIEDOT TYÖPAIKASTANNE

POLIISIL  
RANTAKATU 11  
JOENSUU

OLIVATKO KAIKKI TIEDOT SAMAT VIIKOLLA 11.11.-17.11.1985 ?

KYLLÄ. KÄYTTÄ ENÄ VASTATKO TYÖPAIKKAANNE KOSKEVIIN KYSYMYKSIIN 3-6.  
JOS TIEDOT OVAT OSAKSIKIN MUUTTUNEET TAI VIRHEELLISET, VASTATKAA ENSIN KAIKKIIN KYSYMYKSIIN 3-6.

HUOM! KÄYTTÄ TEHKÖ KORJAUKSIA TÄHÄN!

(B) ALLA ON VÄESTÖKIRJOIHIN MERKITYY AMMATTI-NIMIKKEENNE

NUOREMPI KONSTAPELI

OLIKO AMMATTINNE SAMA VIIKOLLA 11.11.-17.11.1985 ?

KYLLÄ. KÄYTTÄ ENÄ VASTATKO AMMATTIANNE KOSKEVAAN KYSYMYKSEEN 7.  
JOS AMMATTITieto ON MUUTTUNUT TAI VIRHEELLINEN, VASTATKAA MYÖS KYSYMYKSEEN 7.

K Ä Ä N N Ä

## TYÖPAIKKAA KOSKEVAT KYSYMYKSET

<p><b>3</b> Työpaikkanne nimi viikolla 11.11.-17.11.1985</p> <p>Jos olette maanviljelijä tai harjoittatte itsenäistä ammattia, jättäkää rivi tyhjäksi ja ilmoittakaa elinkeinonne kysymyksessä 4.</p> <p><input type="text"/></p>	<p>Ilmoittakaa työpaikan nimi täsmällisesti, kuten esimerkiksi Virtasen autokorjaamo Ky, Valtionrautatiet, Kalevan seurakunta, Tmi Kalle Kuimala ja Oy Kvaritto Ab.</p> <p>Jos työnantajanne on kunta, ilmoittakaa työpaikkanne nimi ilman kunnan nimeä seuraavassa muodossa: Rakennusvirasto, Sosiaali toimisto jne.</p>												
<p><b>4</b> Työpaikan toimiala (ammatinharjoittajan elinkeino)</p> <p>1 <input type="checkbox"/> Maa-, metsä- tai karjatalous. Montako kuukautta teitte metsätöitä 12 viime kuukauden aikana? <input type="text"/> kuukautta</p> <p>2 <input type="checkbox"/> Muu, mikä?</p> <p><input type="text"/></p> <p>(Esim. konepaja, elintarvikemyymälä, peruskoulu)</p>	<p>Toimialalla tarkoitetaan työpaikkanne toiminnan lajia tai tuotantosuuntaa. Kaikilla samassa työpaikassa työskentelevillä on sama toimiala ammatista riippumatta. Toimialoja (elinkeinoja) ovat esimerkiksi</p> <table border="0"> <tr> <td>- hammaslääkärin vastaanotto</td> <td>- siivous</td> </tr> <tr> <td>- lasten perhepäivähoito</td> <td>- puunsahaus</td> </tr> <tr> <td>- viihdemusiikkitoiminta</td> <td>- kiinteistöjen hoito</td> </tr> <tr> <td>- putkiasennustyöt</td> <td>- TV-huolto</td> </tr> <tr> <td>- naisten asustekauppa</td> <td>- tilitoimisto</td> </tr> <tr> <td>- oikeusaputoimisto</td> <td>- kankaiden valmistus</td> </tr> </table> <p>Jos työpaikkassanne harjoitetaan toimintaa eri aloilla, ilmoittakaa niistä pääasiallinen.</p>	- hammaslääkärin vastaanotto	- siivous	- lasten perhepäivähoito	- puunsahaus	- viihdemusiikkitoiminta	- kiinteistöjen hoito	- putkiasennustyöt	- TV-huolto	- naisten asustekauppa	- tilitoimisto	- oikeusaputoimisto	- kankaiden valmistus
- hammaslääkärin vastaanotto	- siivous												
- lasten perhepäivähoito	- puunsahaus												
- viihdemusiikkitoiminta	- kiinteistöjen hoito												
- putkiasennustyöt	- TV-huolto												
- naisten asustekauppa	- tilitoimisto												
- oikeusaputoimisto	- kankaiden valmistus												
<p><b>5</b> Työpaikan sijaintikunta</p> <p>1 <input type="checkbox"/> Sama kuin asuinkunta</p> <p>2 <input type="checkbox"/> Muu kunta, mikä?</p> <p><input type="text"/></p> <p>3 <input type="checkbox"/> Vaihtelee (katso ohjetta)</p> <p style="writing-mode: vertical-rl; transform: rotate(180deg);">Vain yksi vaihtoehto</p>	<p>Asuinkunnaksi katsotaan tässä se kunta, jossa henkilöllä on lainmukainen kotipaikka.</p> <p>Työpaikan sijaintikunta on vaihteleva silloin, kun työtä jatkuvasti tehdään usean kunnan alueella eikä työhön päivittäin saavuta saman toimipaikan kautta (esim. kiitolinjan kuljettajan tai kauppaedustajan työ).</p>												
<p><b>6</b> Työpaikan tarkka lähiosoite</p> <p>1 <input type="checkbox"/> Oman asunnon yhteydessä (esim. maanviljelijällä)</p> <p>2 <input type="checkbox"/> Muu osoite, mikä?</p> <p><input type="text"/></p> <p>(Katuosoite tai sen puuttuessa muu tarkka sijainti)</p> <p>3 <input type="checkbox"/> Työ on liikkuvaa (katso ohjetta)</p>	<p>Ilmoittakaa sen paikan osoite, jossa teitte työtänne.</p> <p>Jos katuosoite puuttuu, antakaa muu tarkka sijainti, kuten esimerkiksi Kaukopään tehtaat, Varuskunta, Mikkolan teollisuusalue tai Kauppatori.</p> <p>Työ on liikkuvaa, jos henkilö päivittäin työssään matkustaa tai liikkuu paikasta toiseen (esim. taksiautoilija tai merimies). Jos työhön kuitenkin säännöllisesti saavutaan saman kiinteän toimipaikan kautta, ilmoitetaan sen osoite (esim. postinkantaja ilmoittaa postitoimipaikan osoitteen).</p>												

## AMMATTIA KOSKEVA KYSYMYS

<p><b>7</b> Ammatti työpaikassa viikolla 11.11.-17.11.1985</p> <p>Ilmoittakaa työtehtäviäenne kuvaava ammattinimike, ei oppiarvoa tai arvonimeä.</p> <p><input type="text"/></p> <p>(Esim. koneinsinööri, atk-suunnittelija, maatalan emäntä)</p>	<p>Ilmoittakaa ammattinne</p> <table border="0"> <tr> <td><b>esim. näin</b></td> <td><b>ei näin</b></td> </tr> <tr> <td>äidinkielenopettaja</td> <td>maisteri</td> </tr> <tr> <td>pankkitoimihenkilö</td> <td>merkonomi</td> </tr> <tr> <td>osastopäällikkö</td> <td>diplomi-insinööri</td> </tr> <tr> <td>kauppias, kampaaja</td> <td>liikkeenharjoittaja</td> </tr> <tr> <td>hitsaaja, raudoitusapumies</td> <td>työmies</td> </tr> <tr> <td>sähköurakoitsija</td> <td>urakoitsija</td> </tr> <tr> <td>putkiasentaja</td> <td>asentaja</td> </tr> <tr> <td>vaiheompelija</td> <td>vaatetustehtaan työntekijä</td> </tr> </table>	<b>esim. näin</b>	<b>ei näin</b>	äidinkielenopettaja	maisteri	pankkitoimihenkilö	merkonomi	osastopäällikkö	diplomi-insinööri	kauppias, kampaaja	liikkeenharjoittaja	hitsaaja, raudoitusapumies	työmies	sähköurakoitsija	urakoitsija	putkiasentaja	asentaja	vaiheompelija	vaatetustehtaan työntekijä
<b>esim. näin</b>	<b>ei näin</b>																		
äidinkielenopettaja	maisteri																		
pankkitoimihenkilö	merkonomi																		
osastopäällikkö	diplomi-insinööri																		
kauppias, kampaaja	liikkeenharjoittaja																		
hitsaaja, raudoitusapumies	työmies																		
sähköurakoitsija	urakoitsija																		
putkiasentaja	asentaja																		
vaiheompelija	vaatetustehtaan työntekijä																		

## KAIKKI VASTAAVAT

<p><b>8</b> Montako kuukautta olitte ansiotyössä tai työttömänä 12 viime kuukauden aikana (18.11.1984-17.11.1985)?</p> <p>Kokopäivätyössä <input type="text"/> kuukautta</p> <p>Osa-aikatyössä vähintään 20 tuntia viikossa <input type="text"/> kuukautta</p> <p>Osa-aikatyössä alle 20 tuntia viikossa <input type="text"/> kuukautta</p> <p>Työttömänä ja etsinyt työtä tai pakkolomalla <input type="text"/> kuukautta</p> <p><input type="checkbox"/> Ei ansiotyössä eikä työttömänä</p>	<p>Työssäoloajaksi lasketaan</p> <ul style="list-style-type: none"> <li>- työ, josta saitte palkkaa tai palkkiota</li> <li>- työ omassa yrityksessä, maatilalla tai itsenäisenä ammatinharjoittajana</li> <li>- ilman varsinaista palkkaa tehty työ perheenjäsenen maatilalla tai yrityksessä</li> <li>- tilapäinen poissaolo työstä sairauden, loman, äitiysloman tai muun sellaisen vuoksi</li> <li>- myös lyhytaikainen ja tilapäinen ansiotyö</li> </ul> <p>Osa-aikatyöksi katsotaan työ, jossa työaika on alle 30 tuntia viikossa tai enintään 120 tuntia kuukaudessa. Jos työaika vaihtelee viikoittain, antakaa tiedot keskimääräisen työajan mukaan (ilman ylityötä). Jos teitte kokopäivätyön ohessa sivutyötä, älkää ilmoittako sitä.</p>
---	---

Yleisohje 1 Tallennetaan kaikki asianmukaisella tavalla annetut lomaketeidot, so. rastit ja vastauskenttien tekstit.

Esim. (1) Lomake voi olla moitteettomasti täytetty tai osin puutteellinen.

(2) Lomakkeella ei esitäytettyä henkilötunnusta lukuunottamatta mahdollisesti ole annettu vastausta yhteenkään kysymykseen.

Poikkeus Tallennuksen yhteydessä poimitaan erilleen manuaaliseen käsittelyyn lomakkeet, joissa

a) henkilötunnus on virheellinen.

b) kaikkia tietoja ei ole annettu niille kuuluvissa kentissä ja lomakkeelta ilmeisesti olisi saatavissa lisäinformaatiota.

Esim. (1) Lomakkeelle tai mukaan liitetyille paperille on kirjoitettu mukaan pitkähkö "kirje".

c) lomake on "käsittämätön"

Esim. (1) Käsiälästä ei saa selvää.

d) lomakkeella on selvästi ristiriitaista tietoa.

Esim. (1) Kysymyksessä 2 (toiminta) rastittu kaikki vaihtoehdot.

e) työpaikka- tai ammattitietojen esitäytön yhteyteen on tehty korjauksia.

Esim. (1) Viivattu yli työpaikan nimi.

(2) Viivattu yli joitakin tiedoista ja annettu vieressä uudet tiedot.

esitäyttötiedot: LINDMAN OY

~~NESTARITIE 4~~

TAMPERE

Kivikuja 5

Näiltä lomakkeilta ei tässä vaiheessa tallenneta tietoja ja niiltä mahdollisesti jo tallennetut tiedot tuhotaan. Lomakkeet toimitetaan Tilastokeskukseen täydennettäväksi ja korjattaviksi, minkä jälkeen ne tulevat uudelleen tallennettaviksi.

Ohje A Tallennuksen yhteydessä suoritetaan yksinkertaista lomaketietojen täydentämistä, kun tietoa ei ole annettu asianmukaisella tavalla, mutta se on annettu muuten selkeästi.

Esim. (1) Ei rastia, mutta kirjoitettu ELÄKELÄINEN.

(2) Tyhjän lomakkeen mukaan on liitetty lappu, jossa lukee OLEN ELÄKKEELLÄ tms.

Ohje B Työpaikan lähiosoitteesta ei tallenneta postinumeroa eikä postitoimipaikkaa, jos tällaiset on ilmoitettu. Jos osoitteeksi kuitenkin on ilmoitettu pelkkä postinumero, postitoimipaikka tai postilokero, tallennetaan kaikki kentässä annettu tieto.

Esim. (1) lähiosoite: KAPULATIE 7 23401 KARKKILA  
tallennetaan: KAPULATIE 7

(2) lähiosoite: PL 374 98700 VALKEALA  
tallennetaan: PL 374 98700 VALKEALA

Ohje C

Selvästi väärässä kentässä annettu tieto pyritään tallentamaan oikeassa kohdassa. Jos siirto ei ole mahdollinen, tallennetaan tieto kuitenkin sen antamiskentässä.

Esim. (1) työp.nimi: MAAJUUSTO OY, PL 67 VALKEAKOSKI  
osoite: MAAJUUSTO OY, LEMPÄÄLÄNTIE 3 VALKEAKOSKI

tallennetaan:

työp.nimi: MAAJUUSTO OY  
osoite: LEMPÄÄLÄNTIE 3

(2) työp.nimi: AUTOMIKKO OY  
ammatti: AUTOMIKKO OY

tallennetaan:

työp.nimi: AUTOMIKKO OY  
ammatti: --

(3) osoite: MAATILAN EMÄNTÄ  
ammatti: --

tallennetaan:

osoite: --  
ammatti: MAATILAN EMÄNTÄ

Ohje D

Jos työpaikan nimeä ei sen pituuden vuoksi voida tallentaa kokonaan, tallennetaan se oleellisin osin käyttäen esimerkiksi lyhenteitä.

Esim. (1) työp.nimi: SEINÄJOEN KANSANTERVEYSTYÖN KUNTAINLIITON  
HAMMASHOITOLA

tallennetaan: SEINÄJOEN KANS.TER.TYÖN KL HAMMASHOIT.

Ohje E

Jos on ilmoitettu tiedot kahdesta työpaikasta, tallennetaan tiedot vain ensimmäisestä.

Ohje F

Lomakkeella ilmoitettua selvästi asiatonta tietoa ei tallenneta. Sen sijaan tiedot, jotka saattavat olla merkityksellisiä joko tallennetaan tai sitten tällaiset lomakkeet poimitaan erilleen.

Esim. (1) työp.nimi: EN TÄYTÄ TÄLLAISIA LOMAKKEITA  
tallennetaan: --

(2) työp. nimi: LOMALLA  
tallennetaan: LOMALLA

(3) osoite: KS.ED.KOHTA  
tallennetaan: KS.ED.KOHTA

Yleisohje 2

Lomake, jonka käsittelyn suhteen tallentaja on epävarma, poimitaan aina erilleen manuaaliseen käsittelyyn.