

SUOMEN METSÄTIETEELLINEN SEURA — FINSKA FORSTSAMFUNDET

ACTA
FORESTALIA FENNICA

63.

ARBEITEN DER
FORSTWISSENSCHAFTLICHEN
GESELLSCHAFT
IN FINNLAND

PUBLICATIONS OF THE
SOCIETY OF FORESTRY
IN FINLAND

PUBLICATIONS DE LA
SOCIÉTÉ FORESTIÈRE
DE FINLANDE

HELSINKI 1956

Suomen Metsätieteellisen Seuran julkaisusarjat:

ACTA FORESTALIA FENNICA. Sisältää etupäässä Suomen metsätaloutta ja sen perusteita käsitteleviä tieteellisiä tutkimuksia. Ilmestyy epäsäännöllisin väliajoin niteinä, joista kukin yleensä käsittää useampia tutkimuksia.

SILVA FENNICA. Sisältää etupäässä Suomen metsätaloutta käsitteleviä kirjoitelmia ja pienehköjä tutkimuksia. Ilmestyy epäsäännöllisin väliajoin.

Finska Forstsamfundets publikationsserier:

ACTA FORESTALIA FENNICA. Innehåller vetenskapliga undersökningar rörande huvudsakligen skogshushållningen i Finland och dess grunder. Banden, vilka icke utkomma periodiskt, omfatta i allmänhet flere avhandlingar.

SILVA FENNICA. Omfattar uppsatser och mindre undersökningar rörande huvudsakligen skogshushållningen i Finland. Utkommer icke periodiskt.

SUOMEN METSÄTIETEELLINEN SEURA — FINSKA FORSTSAMFUNDET

ACTA
FORESTALIA FENNICA

63.

ARBEITEN DER
FORSTWISSENSCHAFTLICHEN
GESELLSCHAFT
IN FINNLAND

PUBLICATIONS OF THE
SOCIETY OF FORESTRY
IN FINLAND

PUBLICATIONS DE LA
SOCIÉTÉ FORESTIÈRE
DE FINLANDE

HELSINKI 1956

Acta Forestalia Fennica 63.

Eino Saari: Alkusanat	IV—VI
Eino Saari: Preface	VII—VIII
1. Lauri Heikinheimo: Metsätyövoiman tutkimusmenetelmä	1—146, 166—182
2. Summary (Method of surveying forest labour)	147—165
Lauri Heikinheimo: Maaseudun miestyövoiman arkiajan käyttö	1—107
Summary (Use of rural manpower in Finland)	108—111
3. Toini Ristimäki: Nuorukaisten ja täysi-ikäisten miesten arkiajan käyttö	1—19
Summary (Annual round of activity of youths and adult men)....	20—21
4. Toini Ristimäki: Kääpiöviljelmien miestyövoiman käyttö	1—96
Summary (Use of manpower of dwarf farms)	97—105
5. Sulo Väänänen: Ammattimaisten metsätyömiesten asunto-olot	1—30
Summary (Housing conditions of professional forest workers in Finland)	31—32
6. Toini Ristimäki, Sulo Väänänen ja Lauri Heikinheimo: Maaseudun elinkeino- ja työttömyysalueet miestyövoiman ajankäytön perusteella ..	1—55, 61—69
Summary (Distribution by industry and unemployment of the manpower in rural districts in Finland)	56—60
7. Lauri Heikinheimo ja Toini Ristimäki: Metsä- ja uittotyövoiman määrä ja rakenne	1—130
Summary (Size and structure of forest and log-floating labour force in Finland)	131—142

Alkusanat

Metsätyötiede, jonka voidaan katsoa Suomessa alkaneen 1920-luvulla, asetti metsätyöllisen entistä paljon tärkeämpään asemaan Suomen metsätieteessä. Tällä nuorella opetus- ja tutkimusalalla metsätyöllistä tarkastellaan ensi sijassa puutavaran tuotannon näkökulmasta, voisipa miltei sanoa tuotantovälineenä. Toisen maailmansodan aikainen yleinen työvoimapula johdatti huomion metsätalouden työllisyyskysymyksiin. Naapurimaassamme Ruotsissa nopea teollistumisvaihe sekä siihen liittyvä ns. pako maaseudulta asutustaajamiin ja siitä johtuva osittainen metsätyövoiman niukkuus olivat omiaan sodan jälkeenkin painottamaan metsätalouden työllisyysongelmia. Ne puolestaan pyrkivät laajenemaan metsätalouden sosiaalipolitiikan muillekin aloille. Tällä tutkimuksen haaralla joudutaan näkökulmaa siirtämään äsken mainitusta. Voidaan sanoa, että siinä tavallaan tarkastellaan metsää työllisten näkökulmasta, kun halutaan saada valaistusta kysymykseen, mikä merkitys metsätaloudella on väestön toimeentulossa. Samalla pyritään selvittämään, minkälaisen elintason metsätyöt pystyvät antamaan sen varassa osittain tai kokonaan elävälle väestön osalle. Tältä taustalta on lähtenyt se julkaisujen sarja, joka nyt Acta forestalia fennican 63:nnessä niteessä seitsemänä erillisenä, mutta toisiinsa liittyvänä tutkimuksena saatetaan julkisuuteen. Eripainoksina osa näitä selvityksiä on jo aiemmin ollut saatavissa.

Eräänä toisen maailmansodan jälkeisenä vuonna keskustelimme virkaveljeni prof. VALTER KELTIKANKAAN kanssa hänen suunnitelmastaan ryhtyä tutkimaan pohjoissuomalaisen metsätyömiehen sielullista asennoitumista eräisiin elämän ilmiöihin. Siitä johduimme Helsingin yliopiston metsätaloustieteellisen laitoksen henkilökunnan keskuudessa pohtimaan tutkimusta, joka selvittelisi metsätyöllisen ja hänen perheensä toimeentuloa sekä metsän osuutta siinä. Tämäkin pohja osoittautui pyrkimyksellemme liian kapeaksi, koska ei ollut saatavissa tietoa, kuinka suuri osa maaseutumme väestöä kuuluu metsätyöllisten suureen joukkoon ja kuinka suuri osa metsätoilla on heidän kokonaistyöpanoksessaan. Joh-

duimme siihen päätelmään, että tämän tapainen tutkimus olisi aloitettava näistä viimeksi mainituista ongelmista, jolloin tehtävä laajeni maaseudun koko työvoiman käsittäväksi.

Tehtävä oli niin suuri, ettei siihen katsottu olevan järkevää ryhtyä ilman valmistavia esitutkimuksia, koska ei mielestämme ollut olemassa tällaiseen työhön valmiiksi kehitettyä tutkimusmenetelmää. Näitä metodeja selvitteleviä töitä suoritettaessa ryhdyttiin neuvotteluihin tilastollisen päätoimiston ja maataloushallituksen kanssa, joista edellinen valmisteli v:n 1950 yleistä väestönlaskentaa ja jälkimmäinen saman aikaista yleistä maatalouslaskentaa. Kytkemällä maaseudun työvoimatutkimus jompaankumpaan näistä tai molempiin oletettiin saatavan aineiston keräys suhteellisen halvaksi ja riittävän laajalle pohjalle. Näiden neuvottelujen tuloksena oli, että tilastollisen päätoimiston ylijohtajan prof. MARTTI KOVERON johdolla työskentelevä, väestönlaskentaa suunnitteleva komitea päätti, että tämän selvittelyn yhteydessä kerätään samalla aineisto Yliopiston metsätaloustieteellisen laitoksen tutkimusta varten. Ja niin saatiin aineisto siihen selvitysten sarjaan, joka käsittää Acta forestalia fennican koko niteen 63. Tälle työlle annettiin nimi maaseudun työvoiman tutkimus.

Tämän laajan tutkimuksen varsinaisena suunnittelijana ja johtajana on toiminut assistenttini maat. ja metsät. toht. LAURI HEIKINHEIMO, jonka käsialaa suuri osa tätä Acta forestalia fennican nidettä on. Hänen lähimpinä apulaisinaan ovat olleet valtiot. kand. TOINI RISTIMÄKI sekä maat. ja metsät. kand. SULO VÄÄNÄNEN. Osa tämän niteen tutkimuksista on heidän kirjoittamiaan ja kulkee heidän nimissään. Osa on yhteistä työtä, ja sen mukaan on tehty merkintä tutkimusten nimilehteen. Tämän tapainen laaja selvittely ymmärrettävästi on luonteeltaan ennen kaikkea ryhmätyötä, jossa kunkin osuutta ei voida tarkalleen rajoittaa. Siihen ovat osallistuneet myös laitokseni muutkin lukuisat tutkimusapulaiset, jotka on erikseen mainittu kunkin tutkimuksen alussa.

Kun työ nyt on saatu päätökseen, haluan lausua kaikille heille kiitokseni heidän sitkeistä ponnistuksistaan yhteisen työn päätökseen viemiseksi.

Koska yliopisto harvoin voi myöntää varoja suuriin tutkimustöihin ja koska Metsätaloustieteellisellä laitoksella tällaisia varoja ei ollut ollenkaan, oli maaseudun työvoiman tutkimus rahoitettava muualta saaduilla varoilla. Muutakin tukea on tarvittu runsaasti yliopiston ulkopuolelta. Metsämiesten Säätiön myöntämällä varoilla saatiin esitutkimukset suoritetuiksi ja päätutkimuksen suunnittelu alulle. Sittenkin eduskunta ja

Suomen Luonnonvarain Tutkimussäätiö ovat myöntäneet varoja, joilla pääosa tutkimusta on suoritettu. Kululaitosten ja yleisten töiden ministeriö, tilastollinen päätoimisto, maataloushallitus ja Suomen Metsätieteellinen Seura ovat niinkään antaneet auliisti tukeaan monessa muodossa. Mieluisa velvollisuuteni on esittää kaikille näille syvä kiitokseni avusta ja ymmärtämyksestä, joita ilman tehtävä olisi jäänyt suorittamatta.

Harras toivomukseni tutkimuksen päättyessä on, että nyt kertaalleen suoritettu perusselvitys ei jäisi ainoaksi, vaan että tämä yhteiskunnallisen tutkimuksen haara saisi vakinaisen aseman jossakin tutkimuslaitoksessa, lähinnä Metsäntutkimuslaitoksessa, jossa puut, metsiköt, metsämaa, metsätyövälineet ja -menetelmät sekä eräät taloustieteelliset aihepiirit jo ovat saaneet vakinaisen edustuksen. Metsätyölliset muodostavat tämän tutkimussarjan mukaan noin 2/3 Suomen maaseudun miestyövoimasta. He ovat varmasti ansainneet pysyvän sijan Suomen metsätieteessä.

Helsingissä heinäkuun 1 päivänä 1956.

Eino Saari

metsätaloustieteen professori

Helsingin yliopiston metsätaloustieteellisen laitoksen esimies

Preface

When forest work studies were started in Finland in the 1920's, the forest worker was given a much more important position in the forest sciences than he had had before. In this new field of education and research, the forest worker is regarded as a factor of production.

The general shortage of labour during the Second World War drew attention to employment problems in forestry. In Sweden the rapid industrialization connected with the so-called «flight» of people from the country to the city areas and the consequent shortage of forest labour emphasised employment problems even after the war.

The interest aroused by the shortage of labour now tends to widen in scope to include sociological and human matters. Thus, the forest worker must be looked at from a viewpoint other than that of the forest work studies. It can be said that when an investigator wants to answer the question of how important forestry is for the population (on both the local and national scale) he must look at the forest from the workers' point of view. Hence, for example, he tries to show what standard of living forest work can give to that part of the population which is partly or entirely dependent on it for its livelihood. From this background stem the seven separate but inter-related papers which are presented to our readers in this the 63rd volume of Acta Forestalia Fennica. Some of them have already been available as reprints.

Shortly after the war, I discussed with my colleague Professor VALTER KELTIKANGAS his intention to start an investigation into the north Finnish lumber jack's attitude to certain aspects of his life. Consequently, we among the staff of the Institute of Forest Economics at the University of Helsinki started to consider an investigation that would throw light on the living conditions of the forest worker and his family, and on the importance of the forest to them. But the scope of the investigation was soon discovered to be too narrow, as it was not known what proportion of Finland's rural population worked in the forest or how important forest work was in their working life. We came to the conclusion that an investigation of this kind had to start with the latter problems. Thus the object of our investigation had to be expanded to include the whole rural labour force.

The task was so big that pilot surveys were considered to be necessary, as in our opinion no previously developed methods were available for this kind of investigation. During these pilot surveys negotiations were started with the Central Statistical Office and the Board of Agriculture, which were respectively preparing a general population census for 1950 and a simultaneous census for agriculture. By connecting

the proposed survey of the rural labour force to one or other, or both, of these censuses, it was presumed that sufficient data could be collected relatively cheaply. As a result of these negotiations, the committee set up to plan the general census (with the Director General of the Central Statistical Office, Professor MARTTI KOVERO as chairman) decided to collect the data required by the Institute of Forest Economics as a part of the general census. It is on these data that the series of papers in this volume are based, under the over-all title of »The Finnish Rural Labour Force Study».

My assistant, Dr. LAURI HEIKINHEIMO, has been chiefly responsible for planning and directing this study, and a big part of this volume of Acta Forestalia Fennica is written by him. His closest co-workers have been Miss TOINI RISTIMÄKI (Master of Political Science), and Mr. SULO VÄÄNÄNEN (Master of Forestry). Some of the papers are written by them under their respective names. The other articles are joint efforts, and the authors responsible are duly named below each title. A big survey of this kind must inevitably be the work of a team which means it is not easy to determine the contribution of each co-worker. Other members of the Institute have also taken part in the work, and their names are mentioned in the preface of each paper. The work being now finished, I want to thank all concerned for their loyal and able labours.

As the University of Helsinki can seldom give money for big investigations, and as our Institute had insufficient funds, the costs of the Finnish Rural Labour Force Study had to be met from other sources. Assistance in various forms has been found from outside the University. The Foresters' Fund paid for the pilot surveys, and the planning of the main survey. Later Parliament and the Fund for Research on Finland's Natural Resources gave the financial aid whereby the main part of the survey was carried out. The Ministry of Communications and Public Works, the Central Statistical Office, the Board of Agriculture and the Society of Forestry in Finland have also given their ready support in many forms. It is my pleasant duty to express my deep gratitude for their help and understanding, without which the task could not have been accomplished.

In conclusion, I hope that this basic investigation will not be the only one in this field in Finland, but that this branch of social research will become a continuing study in one of our research institutions, preferably the Forest Research Institute. In this Institute work is already being done on the biological, technical and some economic aspects of forestry, and it seems most fitting that here too work should be done on the sociological aspects of forestry. According to the papers that follow, forest workers form approximately 2/3 of the male rural labour force in Finland. They certainly deserve a permanent place in Finland's forestry science.

Helsinki, 1st July, 1956.

Eino Saari

Professor in Forest Economics
Head of the Institute of Forest Economics
at the University of Helsinki

METSÄTYÖVOIMAN TUTKIMUSMENETELMÄ

MAASEUDUN TYÖVOIMAN TUTKIMUKSIA I

LAURI HEIKINHEIMO

METHOD OF SURVEYING FOREST LABOUR

FINNISH RURAL LABOUR FORCE STUDIES I

SUMMARY

HELSINKI 1954

Alkusanat

Vuonna 1949 professorit VALTER KELTIKANGAS ja EINO SAARI tekivät Metsämiesten Säätiölle esityksen Suomen metsätyövoiman tutkimisesta. Helsingin yliopiston metsätaloustieteellisen laitoksen assistenttina sain tehtäväkseni tutkimussuunnitelman laatimisen ja mainittuna vuonna aloitettujen esitutkimusten suorittamisen. Esitutkimusten perusteella ja vuoden 1949 metsätalouskomitean, ennen kaikkea professori N. A. OSARAN, aloitteesta laajennettiin suunnitellun päätutkimuksen kohde käsittämään koko maaseudun miestyövoima, vaikka suunnitelman pääpaino edelleenkin oli metsätyövoiman tutkimisessa.

Vuosien 1950 ja 1951 vaihteessa suoritettun yleisen väestönlaskennan yhteydessä kerättiin eduskunnan tarkoitukseen myöntämin varoin mainittua päätutkimusta varten näyte. Aineiston keräystyön hoiti Tilastollinen päätoimisto, mutta muun osan tästä vuoden 1950 maaseudun työvoiman tutkimuksesta sain suoritettavakseni Helsingin yliopiston metsätaloustieteellisessä laitoksessa tarkoitukseen varatun henkilökunnan avustamana. Mainitun tutkimuksen tärkeimmät tulokset julkaistaan MAASEUDUN TYÖVOIMAN TUTKIMUKSIA-nimisenä sarjana, jonka ensimmäisen osan muodostaa tämä esitutkimusten ja päätutkimuksen metodinen tarkastelu. Muut osat sisältävät päätutkimuksen varsinaiset tulokset.

Helsingin yliopiston metsätaloustieteellisen laitoksen esimies professori EINO SAARI on jatkuvasti antamallaan ohjeilla ja neuvoilla vaikuttanut ratkaisevasti tutkimuksen muotoutumiseen, auttanut vaikeimpien pulmien selvittämisessä sekä lukenut käsikirjoituksen tehden varteen otettuja muutosehdotuksia. Olen hänelle kaikesta tästä kiitollinen ja samoin niistä monista neuvotteluista, joita hän aikaansa säästämättä on käynyt hallintoviranomaisten, laitosten ja säätiöiden kanssa saadakseen tutkimukseen tarvittavat varat. Ilman professori Saaren metsätyövoiman ongelmia kohtaan tuntemaa suurta mielenkiintoa ei tutkimusta olisi varmaankaan saatu aikaan.

Vuoden 1950 maaseudun työvoiman tutkimus oli Tilastollisen päätoimiston väestönlaskentaosastolle ja koko sen kenttäorganisaatiolle huo-

Helsinki 1954

Suomalaisen Kirjallisuuden Seuran Kirjapainon Oy.

mattava ylimääräinen rasitus. Mainitun osaston johtajaa filosofian maisteri AARRE TUNKELOA ja hänen alaisiaan pyydän kiittää miellyttävästä yhteistyöstä, monista neuvoista ja ennen kaikkea siitä myötämielisyydestä, jota he, väestönlaskentatoimikunnat ja laskijat ovat osoittaneet tutkimusta kohtaan.

Kulkulaitosten ja yleisten töiden ministeriön työvoima-asiain osaston päällikköä hallitusneuvos JOUNI HAKKARAISTA sekä ylitarkastaja YRJÖ W. LAAKSOSTA kiitän heidän tuestaan tutkimuksen suunnittelussa ja rahoituksessa.

Työni kuluessa olen joutunut kysymään neuvoja paitsi edellä mainituilta, myös monilta muilta eri alojen asiantuntijoilta. Professori HEIKKI WARIS on esitutkimusten suunnittelusta lähtien antanut useita ratkaisevan arvokkaita viitteitä ja lukenut käsikirjoituksen. Tilastomatemaattisten ongelmien ratkaisussa olen saanut apua professori LEO TÖRNQVISTILTÄ, joka myös on tarkastanut käsikirjoituksen, sekä filosofian maisteri VEIKKO HAURULTA. Professorit VALTER KELTIKANGAS ja K. U. PIHKALA ovat etenkin työn suunnittelu- ja esitutkimusten vaiheessa olleet siinä kiinteästi mukana. Hekin ovat tarkastaneet tutkimuksen oikovedosasteella. Maataloushallituksen tilastotoimiston johtajalta maatalous- ja metsätieteiden kandidaatti P. J. SAKSALTA ja hänen alaisiltaan olen työn kuluessa joutunut pyytämään monia erikoistietoja. Kaikille mainituille ja monille nimeltä mainitsemattomille asiantuntijoille esitän kiitokseni.

Metsätaloustieteellisessä laitoksessa suoritetun aineiston työlään käsittelytyön valvomisesta ovat huolehtineet pisimmän ajan valtiotieteiden kandidaatti TOINI RISTIMÄKI sekä maatalous- ja metsätieteiden kandidaatti SULO VÄÄNÄNEN. Myös herra EERO RINNE ja metsänhoitaja T. S. NYMALM ovat osallistuneet työhön miltei sen alusta lähtien. Lausun heille, samoin kuin lähes kahdelle sadalle muulle aineiston käsittelyyn ja esitutkimusten keräystyöhön lyhyemmän aikaa osallistuneelle henkilölle kiitokseni.

Kiitän myös niitä lähes 50 000:a koehenkilöä, jotka eri puolilla maata ovat joutuneet tutkimuksen takia tiedustelujen kohteeksi. Ranuan kunnan koehenkilöt toisessa esitutkimuksessa ansaitsevat erityisen kiitosmaininnan siitä, että he kokonaisen vuoden pitivät kirjaa ajankäytöstään korvauksetta ja yhdenkään sitä keskeyttämättä. Tämä odottamattoman hyvä tulos on luettava myös Ranuan kirjanpidon ohjaajien ansioksi.

Lopuksi mainitsen kiitollisena, että vaimoni maatalous- ja metsätieteiden lisensiaatti ALLI HEIKINHEIMO on monin tavoin minua auttanut työssäni.

Nyt ilmestyvän julkaisun englannin kielisen tekstin on kääntänyt filosofian maisteri HILKKA KONTIOPÄÄ ja tarkastanut Mr. L. A. KEYWORTH, M. A. (Cantab).

Eduskunnan ja Metsämiesten Säätiön lisäksi ovat tutkimusta tukeneet taloudellisesti Suomen Luonnonvarain Tutkimussäätiö, maataloushallitus ja Suomen Metsätieteellinen Seura, joka lisäksi on ottanut tämän tutkimuksen julkaisusarjaansa. Työn viimeistelyvaiheessa olen nauttinut valtion apurahaa. Kaikesta tästä esitän kiitokseni.

Helsingissä, elokuussa 1954.

Lauri Heikinheimo

Sisällysluettelo

	sivu
1. Johdanto	9
11. Terminologia	9
12. Aikaisemmat metsätyövoiman tutkimusmenetelmät	13
13. Otos metsätyövoiman tutkimuksessa	18
14. Ajankäytön määrittäminen työvoiman tutkimuksessa	22
2. Esitutkimukset	27
21. Esitutkimusten järjestely	27
211. I esitutkimus. Haastattelututkimukset Ranuan ja Kontiolahden kunnissa v. 1949	27
212. II esitutkimus. Ranuan kirjanpitoluottamus v. 1950	31
213. III esitutkimus. Ajankäytön haastattelulomakkeiden kokeilu v. 1950	32
22. Esitutkimusten tulokset	33
221. Poimintamenetelmä	33
222. Otoksen suuruus	35
223. Mittausmenetelmä	41
224. Haastattelulomake	51
3. Päättökäytännön tutkimus. Vuoden 1950 maaseudun työvoiman tutkimus	54
31. Päättökäytännön tarkoitus ja rajoitus	54
32. Päättökäytännön menetelmä	57
321. Perusaineiston keräystapa ja suuruus	57
3211. Poimintamenetelmä	57
3212. Mittausmenetelmä	60
3213. Näytteen peittävyys	62
3214. Näytteen suuruus	63
322. Aineiston käsittely	64
3221. Lomakkeiden tarkistus ja kodifointi	64
3222. Reikäkorttiaineistot ja niiden käsittely	65
323. Suurennusmenetelmä	67
3231. Suurennus perusjoukon suuruutta tuntematta	67
3232. Lopullinen suurennusmenetelmä	70
3233. Syntymäpäivämerkintöjen kasautumisilmiö	72
33. Päättökäytännön tulosten luotettavuus	76
331. Poimintaharhat	76
3311. Vastausten kadon aiheuttamat harhat	76
3312. Syntymäpäivämerkintöjen kasautumisen aiheuttamat harhat ..	83
332. Mittausharhat	85
333. Käsittelyerheet	93

334. Satunnaisvaihtelu	94
3341. Työvoiman määrän varianssi	94
3342. Työpanoksen ja muun arkiajan käytön määrän varianssi ..	98
335. Tulokset toisista lähteistä saatujen tietojen valossa	115
3351. Viljelijäperheiden miestyövoiman käyttö	115
3352. Metsä- ja uittotyömiesten lukumäärä	119
3353. Metsä- ja uittotyöpanoksen suuruus ja kausivaihtelu	122
3354. Työttömyysajan määrä ja kausivaihtelu	125
34. Päättökimpuksen menetelmän yleisarvostelu	128
35. Päättökimpuksen kustannukset	131
4. Työpanoksen, työttömyysajan ja työvoiman rakenteen tutkimusmenetelmät saatujen kokemusten valossa	134
Merkkien luettelo — <i>List of symbols</i>	137
Kirjallisuusluettelo — <i>References</i>	142
<i>Summary</i>	147
Liitteet — <i>Appendices</i>	167

1. Johdanto

11. Terminologia

Metsätyövoimasta ja yleensäkin työvoimasta julkaistut esitykset ovat terminologialtaan varsin epäyhtenäisiä. Sen tähden on seuraavaa esitystä varten yritetty rakentaa looginen terminologia, jota riippumatta lähdekirjallisuuden termeistä käytetään tutkimuskohteita selostettaessa.

Kansainvälisen työtoimiston määritelmän mukaan (Employment ... 1948, s. 119) työvoimaan jonakin päivänä kuuluvat kaikki sinä päivänä työssä olevat ja työttömät henkilöt. Työvoiman yksikkönä on henkilö, jota tässä esityksessä nimitetään *työlliseksi*¹ tai *työttömäksi* sen mukaan, onko hän työssä vai työttömänä.

Metsätyöllä tarkoitetaan ruumiillista työtä metsätaloudessa ja jonkin päivän *metsätyöllisillä* kaikkia sinä päivänä pääasiallisesti metsätyötä tekeviä henkilöitä. Koska työttömiä on usein käytännössä vaikea jakaa samalla tavalla alaryhmiin sen työn laadun perusteella, jota asianomaiset pystyisivät ja olisivat halukkaat tekemään, jos työtä olisi tarjolla, *metsätyövoimaan* luetaan seuraavassa vain metsätyölliset.

Jonkin vuoden tai kuukauden *metsätyövoimaan* luetaan kuuluvaksi jokainen henkilö, joka kyseisen vuoden tai kuukauden aikana on ollut vähintään päivän metsätyöllisenä. On huomattava, että näin määritellen vuoden metsätyövoima tulee käytännössä olemaan suurempi kuin minkä tahansa kyseisen vuoden kuukauden tai päivän metsätyövoima. Vastavasti kuukauden eri päivien metsätyövoima on suurimmillaankin tavallisesti pienempi kuin koko kuukauden metsätyövoima. Ero vuoden ja sen työhuipun metsätyöllisten määrissä voi olla hyvinkin suuri. Vuoden, kuukauden tai päivän työvoimassa yleensä tai jonkin erikoisalan työvoimassa on havaittavissa samansuuntaiset erot.

Metsätyövoima (metsätyölliset) on tarpeen jakaa useilla eri perusteilla alaryhmiin. Jos jakoperusteena käytetään työn itsenäisyyttä, jakautuu

¹ Filosofian tohtori MATTI SADENIEMEN kirjoittajan pyynnöstä tekemä ehdotus.

metsätyövoima kahteen osaan, nimittäin *metsätalouden palkkatyövoimaan* eli *metsätyöväkeen*, jonka yksikkönä on *metsätyöntekijä*¹ (ja kun on kysymys vain miehistä *metsätyömies*), sekä *itsenäiseen metsätyövoimaan*, jonka jäsenet ovat *itsenäisiä metsätyöllisiä*. Edelliseen ryhmään kuuluvat työ-sopimussuhteessa olevat metsätyölliset ja jälkimmäiseen kaikki muut metsätyölliset, jotka tekevät metsätyötä omaan lukuunsa tai palkatta perheenjäsenen lukuun.

Vuoden metsätyövoiman määritelmästä johtuu, että metsätyöllisiin jonakin vuonna tulevat sisältymään kaikki päivänkin metsätyötä tehneet, olkoonpa työ tehty omaan tai toisen lukuun. Koska Suomessa valtaosa maaseudun miespuolisesta väestöstä tekee metsätöitä jonkin päivän vuodessa ainakin omassa metsässään, käsite metsätyövoima saa hyvin laajan merkityksen, samoin termi metsätyöväki (HEIKINHEIMO 1954 a). Myös WEIGER (1954, s. 32) antaa käsitteelle metsätyöväki yhtä laajan merkityksen.

Toinen (vuoden) metsätyövoiman (käytännössä pääasiallisesti vain metsätyöväen) jakoperuste on metsätyön merkitys työllisten toimeentulossa. Jos metsätyöllä on työlliselle pääasiallinen merkitys, sanotaan häntä seuraavassa *metsätalouden ammattityölliseksi*, työntekijästä puheellon *ammattimaiseksi metsätyöntekijäksi*. Kaikki muut metsätyölliset ovat *metsätalouden kausityöllisiä* (vastaavasti *metsätalouden kausityöntekijät*). Vastaavat ryhmäkäsitteet ovat *metsätalouden ammatti-* ja *kausityövoima* sekä *ammatti-* ja *kausityöväki*. Kausityöllisellä tarkoitetaan tässä sekä niitä työllisiä, jotka säännöllisesti vuosittain lyhyehköksi ajaksi tulevat metsätöihin, että vain tilapäisesti kerran tai silloin tällöin metsätöitä tekeviä. Koska Suomessa edellinen ryhmä muodostaa metsätalouden kausityöllisten pääosan, on kausi-etuliite tässä laajassa merkityksessä katsottu sopivaksi.

¹ Työntekijä-termi on tässä esityksessä säilytetty samassa merkityksessä, jossa se on Suomen työlainsäädännössä (vrt. esim. SIPILÄ 1938, s. 86 ja VUORIO 1952). Kirjoittajan käsityksen mukaan olisi asiallisempaa, että työlainsäädäntö ottaisi käytäntöön jonkin paremmin asian luonnetta kuvaavan termin, esim. saksan- ja ruotsinkielistä termiä vastaavan *työnottaja*. Silloin työntekijä-termi vapautuisi tarkoittamaan kaikkia työn tekijöitä, joille yllä on jouduttu antamaan uusi termi »työllinen». Vaikka kirjoittaja aikaisemmin (esim. HEIKINHEIMO 1954 a) on käyttänyt työntekijä-käsitettä mainitussa laajassa merkityksessä ja pitää sitä edelleenkin siihen parhaiten sopivana, on tässä ollut pakko luopua siitä turhien sekaannusten välttämiseksi. Työläinen-termi on vastoin kirjoittajan aikaisempia esityksiä kokonaan hylätty, koska työväen järjestöt termin saaman halventavan merkitysvivahteen takia eivät nykyisin pidä sitä suositeltavana.

Kolmas (vuoden) metsätyövoiman (käytännössä pääasiassa vain ammattimaisen metsätyöväen) jakoperuste on työsuhteen vakinaisuus. *Vakinaisella metsätyöllisellä* tarkoitetaan sellaista henkilöä, jonka edellytetään työskentelevän saman omistajan metsätaloudessa jatkuvasti vuodesta toiseen, samaan tapaan kuin esimerkiksi teollisuuslaitoksen pysyvä työvoima (vrt. LÄHDE 1940, s. 80). Kaikki muut metsätyölliset, joita nimitetään *tilapäisiksi*, ryhtyvät metsätyöhön edellyttäen työsuhteen jatkuvan korkeintaan niin kauan kuin kyseinen metsätyömaa on käynnissä. Koska vain käsitteellä vakinainen ammattimetsätyömies on käytännöllinen merkitys, ei sekaannusta synny, vaikka käytetään lyhennettyä nimitystä *vakinainen metsätyömies*.

Neljänneksi metsätyövoima on tarpeen jakaa kolmeen osaan sen perusteella, työskenteleekö se markkinapuun vai kotitarvepuun korjuussa (vrt. SAARI 1922, s. 38 ja PÖNTYNEEN 1936, s. 16) vai muissa metsätöissä. Saamme siten termit *markkinapuun korjuun työvoima*, *kotitarvepuun korjuun työvoima* ja *muu metsätyövoima*.

Metsätyövoima on vielä hyvin usein tarpeen jakaa *hakkuutyövoimaan*, *ajotyövoimaan* ja *muuhun metsätyövoimaan*. Kun kysymys on vain miespuolisesta työvoimasta, hakkuu- ja ajotyöllisen vastineina käytetään termejä *hakkuu-* ja *ajomies*. *Hakkuu-* ja *ajotyömies* tarkoittavat siten työ-sopimussuhteessa olevia hakkuu- ja ajomiehiä.

Paitsi jo mainittuja johdannaisia voidaan edellä selostetuista termeistä johtaa tarpeen vaatiessa lukuisia muitakin, kuten *ammattimainen markkinapuun hakkuutyöväki*.

Metsätyövoimalle ja sen alaryhmille edellä johdetut määritelmät ovat tarpeen mukaan sovellettavissa muidenkin elinkeinohaarojen terminologiaan. Kun Suomessa metsätalous ja uitto kiinteästi liittyvät toisiinsa, joudutaan useimmiten käsittelemään metsä- ja uittotyövoimaa yhdessä. Tällöin on huomattava, että jonkin vuoden metsä- ja uittotyövoimat erikseen mitattuina tavallisesti osittain peittävät toisensa. Vielä suuremmassa määrin menevät päällekkäin maa- ja metsätalouden työvoimat, jos ne mitataan erikseen jonakin vuonna.

Edellä mainittiin, että työvoima kansainvälisen määritelmän mukaan on mitattava henkilöä yksikkönä käyttäen. Jonkin tuotantolaitoksen tai elinkeinohaaran käyttämän työvoiman määrän mittaamiseen henkilö on kuitenkin varsin huonosti soveltuva yksikkö. Erityisesti tämä pitää paikkansa voimakkaasti kausiluonteisissa elinkeinoissa, kuten maa- ja metsätaloudessa. Niinpä esim. KOROLEFF (1951, s. 11) määrittää teollisuuden käyttämän metsätyövoiman (forest labour force) miespäivinä. Tällöin siis

käytetyn työvoiman määrä on mitattu työajan yksiköitten lukumäärällä (vrt. myös HEIKINHEIMO 1951 a). Sekaannusten välttämiseksi on paras käyttää käsitettä »labour input» (esim. JAFFE ja STEWART 1951, s. 205), jonka suomenkielisenä vastineena on ruvettu käyttämään termiä »työpanos» (NIITAMO 1954). *Työpanos* on tuotantoon käytetyn työvoiman määrä työajan yksikköinä mitattuna. Sen mittayksikkönä käytetään työtuntia, -päivää, -viikkoa ja -vuotta, usein myös miestuntia, -päivää, -viikkoa ja -vuotta. Vuosityöntekijä-käsite työvuoden vastineena (esim. Tapaturmatilastoja) on sekaannuksia aiheuttavana syytä hylätä.

Täten määriteltynä työpanos on erittäin käyttökelpoinen käsite nimenomaan metsätalouden työvoimakysymyksissä. *Metsätalouden työpanos* päivinä jonakin vuonna tarkoittaa siten metsätaloudessa sanottuna ajanjaksona käytettyjen työpäivien summaa. Vastaavasti voidaan puhua metsälön tai maatilan työpanoksesta.

Toisaalta voidaan käyttää myös käsitteyhdistelmää *työvoiman työpanos*, jolla tarkoitetaan erään henkilöryhmän työhön käyttämien työtuntien (-päivien jne.) summaa. Myös yksityisen työntekijän työpanos on tarpeellinen käsite, esim. *metsätyöntekijän kokonaistyöpanos* v. 1950 tarkoittaa hänen v. 1950 työhön käyttämiensä päivien summaa. *Metsätyöntekijän metsätyöpanos* sen sijaan tarkoittaa hänen metsätyöhön käyttämiensä päivien lukumäärää, mikä useimmiten on pienempi kuin hänen koko työpanoksensa samana aikana. Omaan tai perheenjäsenen lukuun palkatta tehdyn työn panoksesta käytetään nimitystä *oman työn panos*, työntekijän palkkaa vastaan tekemän työn panoksesta termiä *ansiotyöpanos*.

Arkipäiviksi ja *arkiajaksi* nimitetään seuraavassa kaikkia muita vuoden päiviä paitsi niitä pyhäpäiviä ja muita virallisia vapaapäiviä (vappu ja itsenäisyyspäivä), jotka eivät sisälly työpanokseen. Työpanokseen sisältyviä arkipäiviä nimitetään *työpäiviksi* ja niiden summaa *työajaksi*. Työttömyyteen käytettyjen arkipäivien summa on *työttömyysaika*. *Luppoajalla* tarkoitetaan peitettyyn työttömyyteen käytettyä työaikaa. Tätä käsitettä tarkastellaan lähemmin luvussa 224. On huomattava, että luppoaika sisältyy työpanokseen, päinvastoin kuin *avoimen työttömyyden aika*. Jonkin henkilön *arkiajan käyttö* koostuu toisaalta työpanoksesta alaryhmineen ja toisaalta *muusta arkiajan käytöstä*, kuten sairausajasta, loma-ajasta, avoimen työttömyyden ajasta, opiskeluajasta jne. Vastaavasti puhutaan myös työvoiman arkiajan käytöstä, mikä tarkoittaa henkilöryhmän arkipäivien summaa ja sen jakautumista äsken mainittuihin pää- ja alaryhmiin.

Tilastoterminologiassa on seurattu vähäpätöisin poikkeuksin julkaisua Pohjoismaiden tilastosanasto (1954).

12. Aikaisemmat metsätyövoiman tutkimusmenetelmät

Jos metsätyövoiman tutkimuksiin luetaan vain päätarkoitukseltaan yhteiskuntataloudelliset selvitykset, jää 1900-luvulla nopeasti laajentunut teknillisluonteinen metsätyön tutkimus niiden ulkopuolelle. Viimeksi mainitun pääasiallinen tarkoitus on työmenetelmien ja -välineiden kehittäminen, joten se etupäässä palvelee liiketaloudellisia päämääriä. Kieltämättä raja metsätyövoiman ja metsätyön tutkimuksen välillä on hyvin keinotekoinen, etenkin kun metsätyön tutkimus mitä suurimmassa määrin edistää juuri kansantaloudellisia ja sosiaalipoliittisia pyrkimyksiä kansakunnan elintason kohottamiseksi. Tällainen työvoimatutkimusten aihepiirin rajoitus näyttää kuitenkin hyväksytyin (esim. JAFFE ja STEWART 1951), ja sen tähden sitä käytetään tässäkin.

Tilastolliset selvitykset metsä- ja uittotyövoimasta, joihin seuraavassa yksinomaan rajoitutaan, voivat olla luonteeltaan joko poimintatutkimuksia, jolloin ne perustuvat tavalla tai toisella poimittuun näytteeseen, tai täydellisiä laskentoja, joissa tutkittavan perusjoukon jokaisesta yksiköstä pyritään keräämään tiedot. Toisaalta niissä voidaan kerätä alkutiedot joko työllisiltä itseltään tai muista lähteistä, tavallisimmin työnantajilta, paikallisilta viranomaisilta tai muilta asiaa tuntevilta henkilöiltä. Kolmanneksi voidaan metsä- ja uittotyövoiman tutkimukset sen perusteella, miten tiedot alkulähteestä saadaan numeroin ilmaistuina merkityiksi, siis mittausmenetelmältään, jakaa muistiin perustuviin (haastattelututkimukset ja sellaiset postitiedustelut, joihin annettavat vastaukset perustuvat muistiin) sekä muihin tutkimuksiin.

Kansantaloudelliset metsä- ja uittotyövoiman tutkimukset ovat pyrkineet selvittämään metsätalouden tai metsätalouden ja uiton työpanoksen suuruutta, vaihtelua, tulevaa kehitystä ja toisinaan myös metsätalouden ja uiton (palkka)työvoiman suuruutta, vaihtelua ja työllisyyttä. Yleensä ne ovat ainakin tarkoitettut täydellisiksi laskennoiksi, ja niihin on kerätty tiedot muista lähteistä kuin työllisiltä itseltään.

HELANDER (1923) laski Suomen valtion metsien työpanoksen metsien pinta-alaan ja hakkuumääriin verrattuna vuosina 1911—13 metsähallinnon keräämän palkkausaineiston perusteella. Samanluonteisen tutkimuksen suoritti sittemmin LAKARI (1934) valtionmetsistä ja PELTTARI (1935) yksityismetsistä. Jälkimmäinen oli tosin arvio siitä, kuinka suureksi työvoiman tarve yksityismetsissä voisi määräytyä edellytyksin muodostua, todellisesta yksityismetsien työpanoksesta PELTTARI ei saanut

tietoja. Metsäteollisuuden metsätalouden työpanoksen määriä LINDFORS (1937) on arvioinut hakkuumäärien ja työntuotoslukujen perusteella.

Edellä mainituista poikkesi tutkimusmenetelmältään olennaisesti PÖNTYSEN (1936) tutkimus Suomen markkinapuun hakkuun ja ajon sekä uiton työvoimasta ja työpanoksesta sekä sen kuukausittaisesta vaihtelusta vuosina 1933—34. Se on poimintatutkimus, jonka olosuhteiden pakosta harkintavalintaisen näytteen muodostivat metsähallinnon ja eräiden suurten yhtiöiden työmaaluetteloista kerätyt tiedot metsä- ja uittotyömiesten määristä. Se oli samalla ainakin Suomessa ensimmäinen tutkimus, jossa pyrittiin mittaamaan metsä- ja uittotyövoiman ja sen metsätalouteen ja uittoon antaman työpanoksen kausivaihtelua. PÖNTYSEN selostaa myös varsin yksityiskohtaisesti tämäntapaisten tutkimusten luonnetta ja käytettävissä olevia menetelmiä. Sittemmin myös WICHMANN (1937 ja 1939) Suomen puunjalostusteollisuuden työnantajaliiton jäseniltä kerätyn täydellisen aineiston perusteella selvitti näiden palkkaaman metsä- ja uittotyöväen määrää ja sen kausivaihtelua sekä koostumusta vuosina 1936—38. Alkutiedot WICHMANNIN tutkimuksissa perustunevat etupäässä työmaiden palkkausluetteloihin.

Pohjois-Ruotsin metsätyövoiman suuressa tutkimuksessa (Skogsbygdens ... 1938, ss. 14—16) kysyttiin kaikkien ao. alueen kuntien viranomaisilta mm. metsätalouden ja uiton koko työvoiman (ainoastaan kotitarvehakkuissa työskenteleviä lukuun ottamatta) suuruutta ja sen metsä- ja uittotyöpanosta yhdeltä vuodelta. Alkutiedot perustuivat suurelta osalta ao. viranomaisten muistiin.

Vielä mainittakoon HYRENIUKSEN (1947) arviolaskelmat Ruotsin metsätalouden työvoiman tarpeesta, työvoimataseesta ja sen tulevasta kehityksestä. Tutkimus perustui etupäässä maatalous- ja väestönlaskentojen tuloksiin. Käsillä olevan tutkimuksen kannalta mielenkiintoisempi oli Ruotsin Norrbottenin läänissä talousvuotena 1943/44 suoritettu maatalouslaskenta ja sen yhteydessä tehty tutkimus metsä- ja uittotyöpanoksesta sekä sen kehitysprognosi (WIESLANDER 1948). Mittausmenetelmänä siinä oli kaikista em. tutkimuksista poiketen työllisten vuoden ajankäytön henkilökohtainen haastattelu. Tutkimus oli täydellinen laskenta, haastattelu kohdistui alueen maatalojen kaikkiin täysi-ikäisiin miespuolisiin asukkaisiin, muu alueen metsätalouden ja uiton työpanos arvioitiin veroilmoitusten perusteella. Tähän tutkimukseen palataan jäljempänä useaan otteeseen.

Suomen metsätalouden ja uiton työpanosta selvittelevistä tutkimuksista on vielä mainittava työnantajilta saatuihin, etupäässä palkkaus-

listoihin perustuvat laskennat yhteisuitojen työvoimasta ja työpanoksesta (Uittotilastoa), valtion metsätalouden työpanoksesta (Metsätilastoa), vakuutusvelvollisten metsä- ja uittotyönantajien antamat tiedot käyttämättään työpanoksesta (Tapaturmatilastoa) sekä kulkulaitosten ja yleisten töiden ministeriön kuukausittain metsä- ja uittotyöväen määrästä keräämä, nykyisin täydelliseksi tarkoitettu tilasto (Valtiovarainministeriön kansantalousosaston suhdannesarjat). Lopuksi mainittakoon väestönlaskentojen tilastot ammattimaisen metsä- ja uittotyöväen määrästä ja koostumuksesta useimmissa niissä maissa, joissa tällä ammattiryhmällä on olennainen merkitys.

Pohjoismaiden olosuhteissa metsätalouden työpanoksen mittauksen vaikeus johtuu siitä, että pääosa töistä suoritetaan urakalla ja yksikköpalkalla. Työntekijöiden työaikaa ei sen tähden näissä töissä valvota eikä siitä pidetä kirjaa työnantajan toimesta. Työpanos metsätyömailla jää täten tavallisesti arvioitavaksi joko työtulosten tai palkkasummien avulla. Hakattujen ja ajettujen puutavaramäärien perusteella voidaan työpanos arvioida, jos tunnetaan ns. keskimääräinen työntuotos (tuotoksen ja työpanoksen suhde). Palkkasummista saadaan työpanos arvioiduksi, jos tunnetaan keskimääräinen työaikayksikön ansio, esim. työntekijöiden päiväansioiden aritmeettinen keskiarvo. Edellä selostetuissa työpanostutkimuksissa onkin, Norrbottenin tutkimusta lukuun ottamatta, noudatettu periaatteessa jompaakumpaa näistä menetelmistä. On kuitenkin huomattava, että kummassakin menetelmässä tulos perustuu keskimääräislukuun, jonka suuruus edes tyydyttävällä tarkkuudella on vaikeasti arvioitavissa. Keskimääräinen hakkuutuotos päivässä on riippuvainen mitä moninaisimmista tekijöistä (esim. MAKKONEN 1950), ja siksi sen vaihteluväli suuressa työntekijäjoukossa on laaja. On varsin epäiltävää, voidaanko pienin näyttein, joissa työajan mittaukset ovat mahdollisia, saavuttaa riittävän varmoja tuloksia. Tulosten soveltaminen toiseen ajankohtaan ja toiseen työväkeen, kuin mistä ne alun perin on laskettu, voi helposti johtaa melkoisiin virheisiin.

Keskimääräisen päiväansion arvioimiseksi on myöskin suoritettava työajan määrityksiä. Kun urakkapäiväansiot Suomessa on asetuksella säännöstelty, on aina olemassa vaara, että näitä sellaisinaan käytetään työpanosta arvioitaessa (vrt. HEIKINHEIMO 1951 b). Metsätöiden päiväansiotilastoja ei liene saatu tyydyttävälle kannalle missään maassa, jossa vallitsee etupäässä yksikköpalkkajärjestelmä. Koska sekä työntuotoksen että päiväansion keskiarvon arviot yleensä tarkoittavat kohtalaisen hyvissä olosuhteissa (esim. ammattitaitoisen miehen 8-tuntisena työpäivänä)

saavutettuja tuloksia, ovat täten lasketut työpanosluvut helposti liian pieniä.

Työnantajilta kerättyjen työpanostietojen toinen periaatteellinen puutteellisuus on siinä, että tällaiset tilastot on miltei mahdoton saada peittämään koko kyseinen työvoima (JAFJE ja STEWART 1951, s. 49). Varsinkin metsänomistajaperheiden jäsenten oman metsätyön panosta on vaikea täten arvioida. Täydellisten laskentojen haittana on vielä mainittava niiden monet vaikeudet ja virhelähteet, esim. tilastoväsymysilmiö.

Metsätyöväen suuruus jonakin määrähetkenä, esim. määrättyinä päivinä, on epäilemättä helpommin mitattavissa kuin sen metsätyöpanos. Kun sen suuruus on tunnettu, on ajateltavissa, että lyhyehkön ajanjakson työpanos lasketaan kertomalla työväen määrä työntekijän keskimääräisellä metsätyöpanoksella sanottuna ajanjaksona (PÖNTYNEN 1936, s. 11, vrt. myös luku 3353). Siinäkin virhemahdollisuudet ovat suuret. Metsätyöllisten täydellisessä tilastoinnissa on luonnollisesti samat vaikeudet kuin edellä jo mainittiin.

Väestönlaskentojen mahdollisuus mitata ammattimaisen metsä- ja uittotyöväen suuruus on etupäässä riippuvainen siitä, kykenevätkö asianomaiset henkilöt itse päättämään, kuuluvatko he mainittuun työväen ryhmään. Tutkijoittekin on ollut varsin vaikea vetää raja tämän työväen ja sitä useimmissa metsätalousmaissa paljon lukuisamman kausityöväen välille, jolle metsä- ja uittotyöt ovat vain sivuammatti (esim. TREITSCHKE 1928, s. 14, SAUER 1929, ss. 3—6 ja WEIGER 1954, ss. 30—32). Sitä paitsi on todettu, että ammattinimitystä »metsätyöntekijä» kohtaan voidaan tuntea vastenmielisyyttä sen vähäisen arvostuksen takia (esim. CARLSSON 1950, s. 88). Niinpä Suomen ensimmäisessä yleisessä väestönlaskennassa v. 1950 tiedusteltiin metsätalouden ja uiton ansiotyöpanoksen suuruutta mainittuna vuonna ja kyseinen raja pyrittiin vetämään sen perusteella. Tätä menetelmää on selostettu luvussa 3352.

Metsä- ja uittotyövoiman sosiaalipoliittiset tutkimukset ovat pääasiassa olleet poimintatutkimuksia. Tämä on ymmärrettävissä, koska metsä- ja uittotyöväen asunto-oloja, päivän tai vuoden käyttöä, tuloja, ravintoa, perhesuhteita, mielipiteitä, jopa elämäkertaakin koskevista selvityksissä on usein mahdoton ajatella koko perusjoukon tutkimista. Vain joitakin yleispiirteisiä tietoja metsä- ja uittotyöväen työskentelyolosuhteista ja elintavoista (SNELLMAN 1914, s. 183) tai työllisyydestä (Skogsarbetarnas ... 1916, s. * 51) on voitu suorittaa laajamittaisina täydellisinä laskentoina edellisessä tapauksessa Suomen metsätyönantajille, jälkimmäisessä tapauksessa Pohjois-Ruotsin kuntiin lähetettyjen tiedustelujen

pohjalla. Lukuisat Saksan valtionmetsien työväen oloja koskevat tutkimukset, kuten jo mainitut TREITSCHKEN, SAUERIN ja WEIGERIN julkaisut, perustuvat hoitoalueista kerättyihin täydellisiin tilastoihin. Samoin on muitakin pienekköjä alueita voitu tutkia näytettä ottamatta. Mainittakoon kolmea Ruotsin kuntaa koskeva WALLANDERIN (1948) tutkimus metsätyöväen tuloista, työpanoksesta, iästä, muuttoliikkeestä ja perhesuhteista. WALLANDERIN joutui pääaineistonsa keräämään näytteenä selvitellessään metsästä paon syitä perheiden haastattelututkimuksella (mt. s. 279).

Mainitunlaatuissa tutkimuksissa on useimmiten käytetty harkintavalintaa. WALLANDERILLA se kuitenkin oli tasavälinen. Niinpä Pohjois-Ruotsin ensimmäisessä laajassa metsätyöväen tutkimuksessa (Skogsarbetarnas ... 1916) pääaineisto metsätyöväen olosuhteista työmailla kerättiin paikallisviranomaisien harkinnanvaraisesti määräämistä »edustavista» työmaista (mt. s. *53). Samaan tapaan suoritettiin tutkimuskohteiden valinta vastaavassa Suomen tutkimuksessa (Tutkimus ... 1923, s. 1). Ruotsalaiseen tutkimukseen liittyi vielä selvitys vakinaisen metsätyöväen elinehdoista. Siihen antoivat aineiston ne maatilan omistajat, joilla oli mainittua työväkeä ja jotka olivat halukkaat antamaan tietoja (mt. s. *47).

Erikoislaatuinen näyte tutkittiin Pohjois-Ruotsin toisessa laajassa metsätyöväen tutkimuksessa (Skogsbygdens ... 1938, s. 210). Tapaturmavakuutettujen ja vahingoittuneiden metsä- ja uittotyömiesten työvuoden käytöstä ja vuosituloista vahingoittumisen jälkeen kyselemällä kerätyt tiedot muodostivat ko. tutkimuksen erään osan aineiston.

Norjassa suorittivat ØDEGÅRD ja RESEN-MANDT (1939) v. 1937/38 tutkimuksen metsätyöväen koostumuksesta ja vuosituloista käyttäen näyttekysikköinä niitä metsä- ja maataloustyöväen liittoon kuuluneita metsätyöntekijöitä, jotka halusivat antaa tietoja. Suomen sosiaaliministeriön metsätyöväen työmaa-asunnoista suorittama vajaanmuotoinen poimintatutkimus (LINNOILA 1939, s. 485) tehtiin keräämällä tiedot rajoitetun alueen kaikista määräsuuruuksista metsätyömaista. CARLSSON (1950, s. 88) poimi näytteensä Ruotsin Norlannissa henkikirjojen perusteella siten, että paikalliset asiantuntijat ilmoittivat, ketkä henkikirjaan merkityistä olivat metsätyömiehiä. CARLSSON tutki metsätyömiesten lukumäärää, ikäjakautumista ja asennoitumista työhön.

KANTOLA (1949, s. 69) tutki Suomen metsätyöväen ikäjakautumista ja palkkatuloja näytteestä, johon kuului joukko mielivaltaisesti valittuja työmaita. Samantapaisia lienevät olleet WUORISTON (1934) tutkimukset

ikäjakautumisesta ym. Suomen virallinen metsätyöväen päivänsiotilasto saadaan näytteestä, johon ennen kaikkea valituksia aiheuttaneet työmaat kuuluvat. Muihin metsätyöpalkkatilastoihin, joista osa on myös täydellisiä, ei tässä ole syytä puuttua.

Yleisenä piirteenä edellä luetelluille poimintatutkimuksille on näytteen harkintavalinta sekä se, että ne kohdistuvat metsätyömaihin. Tavallisesti tiedot on kerätty haastattelemalla ao. työmaiden työntekijöitä tai siten, että havainnoitsija on kirjoittanut tekemänsä omakohtaiset huomiot lomakkeeseen. Usein lienee työmaan työnjohtaja ollut merkittävästi apuna lomakkeiden täyttämässä.

Harkintavalintaiset, mielivaltaiset ja vajaamuotoiset näytteet otanta-tutkimuksen teoria on hylännyt harhan suuren vaaran ja satunnaisvaihtelun mittaamisen vaikeuden tähden. Työmaan käyttäminen näytetyksikkönä on tietysti asianmukaista silloin, kun tarkoitus on nimenomaan tutkia *työmaiden* asunto-, ruokailu- ym. olosuhteita. Sen sijaan metsä- ja uittotyöväen ajankäytön, vuositulojen, yhteiskuntaryhmyksen, ikäjakautuman yms. seikkojen tutkiminen sanotulla tavalla voi johtaa varsin harhaisiin tuloksiin. On nimittäin huomattava, että työvoiman koostumus työmailla ainakin Pohjoismaissa vaihtelee eri vuodenaikoina. Siten tutkimuksen ajankohta voi vaikuttaa huomattavasti sen tulokseen, joka niin ollen ei ehkä anna luotettavaa kuvaa edes ko. vuoden metsä- ja uittotyöväen sosiaalisista oloista. Niinpä aiemmin mainituista tutkimuksista CARLSSONIN, WALLANDERIN sekä ÖDEGÅRDIN ja RESEN-MANDTIN näytteissä yksikkönä oli henkilö, samoin Ruotsin tapaturma-aineistossa. Henkilön käyttäminen näytteen poimintayksikkönä on sitäkin suositeltavampaa, kun metsä- ja uittotyöllisistä varsin huomattava osa työskentelee pienillä työmailla, joiden saaminen näytteeseen voi tuottaa miltei voittamattomia vaikeuksia, jos työmaa on poimintayksikkönä.

13. Otos metsätyövoiman tutkimuksessa

Näytteen poimintamenetelmistä suositeltavimpana tilastotiede pitää todennäköisyysvalintaa, jolla saatua näytettä nimitetään *otokseksi*. Edellä on selostettu, miksi metsätyövoiman tutkimuksessa otoksen poimintayksikkönä henkilö on useimmiten suositeltavampi kuin työmaa. Sellainenkaan kaksiasteinen poiminta ei ole suositeltava, jossa ensin poimitaan näytteeseen joukko työmaita ja sitten näillä työskentelevästä työvoimasta joukko henkilöitä. Poikkeuksen muodostavat tutkimukset, jotka kohdis-

tuvat nimenomaan määräsuuruisten työmaiden olosuhteisiin eikä koko metsä- ja uittotyövoimaan tai metsä- ja uittotyöväkeen jollakin alueella. Sen sijaan voidaan hyvin ajatella, että kaksiasteisessa poiminnassa ensin valitaan joukko pienalueita, esim. kuntia, valtakunnasta ja niistä poimitaan jokin määrä henkilöitä. Kolmiasteinen poiminta, jossa poimintajärjestys on pienalueet, työmaat ja henkilöt, voi myös olla käyttökelpoinen esim. sellaisessa tapauksessa, että on tutkittava työmaiden laadun vaikutusta työntekijöiden olosuhteisiin.

Henkilöä poimintayksikkönä (ilman työmaita esipoiminta-asteena) käytettäessä on suurin vaikeus *perusjoukon rajoittamisessa*. Jonkin alueen metsä- ja uittotyöllisistä tässä tarkoitettussa mielessä on perin harvoin saatavissa luetteloita. Suomen ja Ruotsin väestöluetteloita, joihin on merkitty pääammatti, ei edellä selostetuista syistä voida pitää edes ammattimaisen metsä- ja uittotyöväen luotettavana luettelona. CARLSSON (1950, s. 88) tosin antoi paikallisten asiantuntijoiden merkitä sellaiseen luetteloon metsätyömiehet, mutta tätäkin tapaa on pidettävä jossain määrin epävarmana. Kausityömiehet, joilla Pohjoismaissa on suurempi merkitys kuin ammattimaisilla metsätyömiehillä, jäävät näet helposti merkitsemättä. Joissakin tapauksissa tämä menetelmä taloudellisuutensa takia voi olla suositeltava. Kaikkien metsä- ja uittotyöllisten luettelointi mainitulla tavalla on tuskin ajateltavissa.

Eräitä mahdollisuuksia on kuitenkin olemassa metsä- ja uittotyövoiman sisältämän perusjoukon rajoituksiin. Pohjoismaiden olosuhteissa kaupunkien ja kauppaloitten väestö varsin harvoin tekee metsä- ja uittotyötä. Naiset niihin eivät sanottavasti osallistu (SAARI 1928, ss. 285—287) eivätkä kansakouluiässä olevat pojat (vrt. kuitenkin lukua 31). Saksassa sen sijaan on naisten ja lasten käyttäminen metsätöissä verrattain yleistä (WEIGER 1954, s. 44).

Esimerkiksi Suomessa ja Pohjois-Ruotsissa, jossa hyvin huomattava osa maaseudun (maaseudulla tarkoitetaan seuraavassa maalaiskuntia) miespuolisesta kansakouluiän sivuuttaneesta väestöstä ainakin jossain määrin osallistuu metsä- ja uittotöihin, voi olla edullisinta poimia metsä- ja uittotyövoiman tutkimuksia varten otos koko tästä väestöstä. Tällöin saavutetaan mm. se tärkeä etu, että näyte peittää koko metsä- ja uittotyövoiman eikä ainoastaan metsä- ja uittotyöväkeä. Kun huomattava ja suureneva määrä metsän hakkuu- ja ajotöistä suoritetaan metsänomistajien toimesta ja heidän perheenjäsentensä työvoimalla, on raja metsätyöväen ja itsenäisen metsätyövoiman välillä varsin epäselvä, vaihteleva ja kansantaloudellisesti usein vähämerkityksinen. Tiedot yksinomaan

metsätalouden ansiotyöpanoksesta ja sen muutoksista voivat antaa suorastaan harhaanjohtavan kuvan, jos (niinkuin valitettavan usein tehdään) katsotaan sen osoittavan myös koko metsätyöpanoksen suuruutta ja vaihtelua.

Toinen, vielä tärkeämpi näkökohta, joka puoltaa metsä- ja uittotyövoimatutkimuksen otoksen laajentamista koko mainitun perusjoukon peittäväksi, on metsä- ja uittotöiden ja muiden maaseudun kausiluonteisten töiden moninainen yhteenkytkeytyminen. Useimmissa metsätaloutta harjoittavissa maissa, esim. Pohjoismaissa, Saksassa, Itävallassa, Puolassa, Kanadassa ja USA:ssa sama työvoima, joka osallistuu vuosittain metsä- (ja uitto)töihin, tekee myös maataloustöitä, vieläpä huomattavasti enemmän kuin metsätöitä (em. tutkimukset metsätyöväen sosiaalisista oloista sekä ENDER 1950, KLIMEK 1950, s. 282 ja MACDONALD 1950, ss. 292—244). Näin ollen maa- ja metsätalouden työvoimakysymykset kytkeytyvät niin läheisesti yhteen, että niitä erillisinä on vaikea perusteellisesti käsitellä (vrt. myös BERESFORD-PEIRSE 1950, ss. 221—222 ja OSARA 1950). Suomen ja Ruotsin olosuhteissa sama työvoima vielä suorittaa osan maantien-, rautatien- ja talonrakennustöistä.

Mainittakoon vielä kolmantena näytteen laajaa peittävyyttä puoltavana näkökohtana, että maaseudun koko miestyövoiman tutkiminen sellaisenaan on tärkeää. Kun seuraavassa todetaan Suomessa valtaosan maaseudun miespuolisesta työvoimasta kuuluvan juuri metsä- ja uittotyövoimaan, voidaan hyvin puoltaa sellaista menettelyä, että tiedot maaseudun koko miespuolisesta työvoimasta saadaan samalla kertaa metsä- ja uittotyövoimatutkimuksen kanssa.

Näin ollen sellaisissa metsä- ja uittotyövoiman tutkimuksissa, joiden päätarkoitus on kansantaloustieteellinen, voidaan Suomen kaltaisissa olosuhteissa muun yleisluonteisen työvoiman tutkimuksen puuttuessa käyttää edellä tarkoitettua, peittävyydeltään laajaa otosta. Sosiaalipoliittisissa metsä- ja uittotyövoiman tutkimuksissa, joissa päähuomio kohdistuu ammattimaiseen palkkatyövoimaan, tällainen otos on liian laaja. Ammattimaisen metsätalouden ja uiton palkkatyövoiman rajoittaminen erillistutkimuksessa tuottaa kuitenkin suuria vaikeuksia. Yhdistämällä kansantaloustieteellinen ja sosiaalipoliittinen metsä- ja uittotyövoiman tutkimus saadaan myös mainitusta työntekijäryhmästä otos edellä mainitun otoksen alaryhmänä.

Kun metsätyövoiman tutkimuksesta täten tulee luonteeltaan yleinen maaseudun miestyövoiman tutkimus, on syytä tarkastella, mihin perusjoukkoon yleiset työvoiman tutkimukset kohdistuvat ja minkälaista poimintamenetelmää niihin on sovellettu.

Työvoiman määrällä voidaan tarkoittaa useita eri tavoin mitattuja suureita (JAFJE ja STEWART 1951, ss. 13—47). Siihen ei yleensä lueta esim. kotonaan kotitaloudessa työskenteleviä perheenjäseniä. Työvoima tässä merkityksessä on mitattavissa joko niin, että siihen sisällytetään ns. ammatissa toimiva väestö (gainful workers) tai työvoima (labour force) siinä mielessä kuin se käsitetään mm. USA:ssa ja Kanadassa kuukausittain suoritettavissa (otokseen perustuvissa) työvoiman arvioinneissa. Ammatissa toimivan väestön määrä mitataan yleensä väestönlaskennoissa ja siihen sisällytetään silloin ne henkilöt, jotka laskentahetkellä ilmoittavat tavallisesti työskentelevänsä yhdessä tai useammassa ammatissa. Oma kotitaloustyötä ei tällöin pidetä ammattina. Kyseiseen työvoimaan eivät näin ollen tule luetuiksi esim. sellaiset henkilöt, jotka vasta hakevat ensimmäistä työpaikkaansa laskentahetkellä, usein eivät myöskään tilapäisesti työssä olevat. Ylipäänsä jätetään henkilön itsensä ratkaistavaksi, kuuluuko hän työvoimaan vai ei.

Ammatissa toimivan väestön määrä antaa monessa suhteessa puutteelliset tiedot työvoimasta. Ensinnäkään se ei kykene ilmaisemaan työvoimaan sisällytettävän työttömyyden määrää. Toisaalta erityisesti kausiluonteisissa elinkeinoissa, kuten maa- ja metsätaloudessa, työvoiman määrä tällä tavoin ilmaistuna saa varsin epämääräisen sisällön. Niinpä käsitteeseen maatalouden ammatissa toimiva väestö sisältyy Suomessa pääosa metsätalouden ja hyvin huomattava osa talonrakennusteollisuuden, maantien- ja rautatienrakennuksen, jopa käsi- ja pienteollisuudenkin sekä liikenteen työvoimasta.

USA:n edellä mainitussa otantatutkimuksessa luetaan työvoimaan kaikki otokseen poimitut 14-vuotiaat tai sitä vanhemmat henkilöt, jotka ovat tutkimusviikon aikana olleet työssä tai hakeneet työtä (JAFJE ja STEWART mt. ss. 40—47). Koehenkilöiltä pyydetään joukko tietoja heidän ajankäytöstään tutkimusviikolla, ja näiden vastausten perusteella tarkan kaavan mukaan päätellään, mihin työvoiman osaryhmään he kuuluvat, vai luetaanko heidät työvoimaan kuulumattomien ryhmään. Myös palkattomat, yrityksen haltijaa avustavat perheenjäsenet, jotka viikon aikana ovat olleet tässä työssä vähintään 15 tuntia (palkatonta kotitaloustyötä lukuun ottamatta) katsotaan työvoimaan kuuluviksi. Myöskin USA:n väestönlaskennassa tehdään nykyisin samanlaisen kaavan mukainen työvoiman määrän mittaus aikaisemman ammatissa toimivan väestön mittauksen sijaan.

USA:ssa kehitetty työvoiman arvioimismenetelmä antaa täten tutkimusviikon työvoimatilanteesta paljon selvemmän ja väestönlaskennoissa myös yksityiskohtaisemman kuvan kuin ammatissa toimivan väestön

arvioinnit. Varsinkin työllisyystilanne tulee valaistuksi myös suhdanne-tutkimuksia tyydyttävällä varmuudella ja nopeudella, kun kuukausittaiset estimaatit voidaan julkaista jopa 30 päivää aineiston kokoamisen jälkeen. Mainitun tutkimuksen otos käsittää USA:ssa n. 25 000 ruokakuntaa, jotka valitaan liittovaltion alueelta moniasteisella HANSENIN ja HURWITZIN kehittämällä aluepoiminnalla (esim. WEISS ja WILLIAMS 1952 sekä STEINBERG 1953). Perusjoukon muodostavat liittovaltion kaikki ruokakunnat. Myös kansainvälinen työtoimisto suosittelee työvoiman mittauksissa erityistä, sitä varten kehitettyä väestön otantatutkimusta (Employment ... 1948, s. 125).

Vaikka USA:n työvoimatutkimuksen poimintamenetelmää pidetään väestöstä poimitun otoksen parhaimpana esimerkkinä (esim. DALENIUS 1949, s. 32), ja sitä on sovellettu useissa muissakin maissa, ei siihen ole tässä tarpeen enempää puuttua. Sen sijaan tarkasteltakoon mainituissa tutkimuksissa käytettyä mittausmenetelmää, joka on luonteeltaan ajankäytön haastattelu.

14. Ajankäytön määrittäminen työvoiman tutkimuksessa

USA:n työvoiman tutkimuksissa ja muiden maiden vastaavissa soveltuksissa koehenkilöiltä kysytään, onko havaintoviikko pääasiallisesti käytetty työhön, työn hakemiseen, kotitalouteen, koulunkäyntiin vai muuhun. Jos vastaaja lukeutuu ensimmäiseen ryhmään, kysytään häneltä myös työpanoksen kokonaismäärää tunneissa sekä sen pääasiallista käyttöä (ammatti, työpaikan laatu, työsuhde). Näiden sekä eräiden lisätietojen perusteella luokitellaan koehenkilöt työllisiin elinkeinohaaroittain, työttömiin ja työvoimaan kuulumattomiin. Työpanoksen määrä vaikuttaa tähän luokitukseen vain sikäli, että sen täytyy olla vähintään 15 tuntia, ennen kuin palkatta yrittäjää avustanut perheenjäsen luetaan työlliseksi (JAFFE ja STEWART mt. ss. 40—43).

Vaikka USA:n työvoimatutkimuksessa suoritetaan eräänlainen, tosin epätäydellinen, ajankäytön tiedustelu, ei työvoiman kokonaistyöpanokseen enempää kuin muihinkaan ajankäytön kokonaismääriin pyritä.

Tämä mittausmenetelmä ei kuitenkaan sovellu sellaisenaan käytettäväksi vähän teollistuneissa maissa. Niinpä JAFFE ja STEWART (mt. ss. 469—476) arvostellessaan yrityksiä soveltaa amerikkalaista työvoiman otantatutkimuksen menetelmää Puerto Ricossa ja Japanissa päätyvät tulokseen, että näiden tilastojen käytännöllinen arvo on »hyvin pieni».

Kyseinen mittausmenetelmä on nimittäin kehitetty sellaisia olosuhteita varten, joissa työssä oleva henkilö todella tekee tehokasta työtä koko viikon ajan, ts. on välttämätön kyseiselle tuotannolle. Milloin hän taas ei ole työssä, vaan hakee työpaikkaa, on hän selvästi työtön. Laskusuhdanteissa tilasto osoittaa työvoiman pienenevän ja työttömyyden suurenevan.

Sen sijaan Japanissa, jossa luppoaika on luonteenomainen ilmiö, samalla menetelmällä kerätty tilasto laskusuhdanteissa osoittaa työvoiman suurenevan ja työttömyyden pyrkivän kaikissa olosuhteissa pysyttelemään väistämättömän kitkatyöttömyyden tasolla, siis hyvin vähäisenä. Tuotoksen kokonaismäärää on pakko varsinkin maanviljelysseuduilla pyrkiä lisäämään välittämättä siitä, minkälaiseksi työstä saatava tulo ja muut työolosuhteet muodostuvat. Työttömyys peittyi vähätuottoiseen tai peräti tuottamattomaan työhön ja voi siten olla huomattavan suuri, vaikkei se tilastoissa näy.

Mainituista syistä on vielä suhteellisen vähän teollistuneissa maissa, joissa varsinkin maataloudella on työvoiman käytössä huomattavasti suurempi merkitys kuin USA:ssa, työvoiman arviointimenetelmät sovellettava asianomaisen maan olosuhteisiin ja tarpeisiin. JAFFE ja STEWART (mt. s. 54) suosittelevat työvoiman arvioimista työpanoksen määrää perusteena käyttäen. Joissakin tapauksissa heidän mielestään voisi olla tarpeen antaa työntuotoksen tai työstä saatavan tulon määrätä, milloin joku henkilö on luettava työlliseksi, jotta peitetysti työttömät voitaisiin erottaa työllisistä. Vastaavasti on työttömyyden määrää arvioitaessa käytettävä toisenlaisia menetelmiä kuin USA:ssa. Sielläkin kysymys omaa työtään tekevän maatalousväen työttömyyden ja työllisyyden määrittämisestä on tuottanut vaikeuksia (DUCOFF ja BANCROFT 1945), mutta kyseisen ryhmän suhteellinen pienuus ei ole pakottanut muuttamaan mittausmenetelmää ainakaan toistaiseksi. USA:ssa henkilöä ei voida merkitä työttömäksi, jollei hän ole aktiivisesti etsinyt työpaikkaa, olkoon hänen työpanoksensa tutkimusviikon aikana ollut kuinka pieni hyvänsä, ja vaikka hän olisi jäänyt omaan työhönsä vain sen tähden, ettei sopivaa ansio-työtä ole saatavissa.

Suomessa on, varsinkin maaseudulla, havaittavissa periaatteeltaan samantapainen vajaatyöllisyyden ja peitetyn työttömyyden ilmiö kuin Japanissa. Erityisesti toisen maailmansodan jälkeen tähän seikkaan on ruvettu kiinnittämään huomiota (esim. HEIKINHEIMO 1950 a, KAHMA 1953 ja SALONEN 1953).

USA:n työvoiman otantatutkimuksen muuntaminen sellaiseksi, että työn tuotos tai tuotto määräisi rajan työssä olevan ja peitetysti työttö-

män koehenkilön välillä, tuntuu varsin vaikealta toteuttaa. Ajankäytön mittaamenetelmä siinä voidaan sen sijaan kehittää täydellisemmäksi niin, että työvoiman koko arki-aika ja sen jakautuminen eri elinkeinohaarojen työpanoksiin sekä muihin ajankäyttöryhmiin saadaan arvioiduksi esim. vuoden ajalta. Jos vastaavien elinkeinohaarojen tuotos on tunnettu, on näitä lukuja vertaamalla mahdollista arvioida, kuinka tehokkaasti työpanos on tullut käytetyksi. Mahdollisesti voidaan täten laskea myös luppoajan määrä, jolle sitä saada suoraan koehenkilöltä mitatuksi.

Ajankäytön mittaamenetelmiä on kehitetty varsinkin työtieteessä, sosiologiassa ja kodin taloustieteessä. Niiden metodiikasta antaa mm. SAURIO (1947, ss. 11—17) yleiskatsauksen. Hän toteaa, että koehenkilöiden lukumäärä on kääntäen verrannollinen mittauksen tarkkuuteen ja havaintoajan pituuteen. Tähän voidaan lisätä, että aineiston laajuus on ennen kaikkea riippuvainen käytävissä olevista varoista, mutta myös näytteen poimintamenetelmästä, aineiston heterogeenisyydestä sekä tietojen keräystavasta.

Työtieteen ns. työaikatutkimuksissa koehenkilön toimintoja seurataan välittömästi kellon avulla (SÄLLFORS 1945, s. 121). Samaa menetelmää on sovellettu myös kodin taloustieteessä (SAURIO mt., ss. 11—17). Tällöin koehenkilöiden määrä on ollut pakko supistaa muutamiin harvoin. Tämän ajankäytön mittaustavan soveltamista työvoimatutkimukseen yleensä ja varsinkin otokseen voidaan tuskin pitää kannattavana.

Huomattaviin mittaustarkkuuksiin ja yksityiskohtaiseen ajankäytön jaotteluun on päästy myös kirjanpitomenetelmällä, jossa koehenkilöt itse suorittavat ajankäyttönsä muistiin merkitsemisen välittömästi toimintojen yhteydessä. Milloin koehenkilöt on valittu vapaaehtoisesti ilmoittautuneista (SAURIO mt. ss. 18—20), voidaan koehenkilöiden määrä saada verraten suureksi, jopa tuhansiksi.

Todennäköisyysvalintaa on tällaisissa tutkimuksissa kuitenkin hyvin harvoin sovellettu. Niinpä kirjanpitotiloilla suoritettuja Suomen maatalouden kannattavaisuustutkimuksia, joita osittain on pidettävä tähän ryhmään kuuluvina ajankäytön tutkimuksina, juuri tämän tähden pidetään tuloksiltaan harhaisina (SAULI 1951, s. 57). Osastopäällikkö FOLKE LARSSONIN Jordbrukets utredningsinstitutista tekijälle antamien tietojen mukaan kokeillaan Etelä-Ruotsissa parhaillaan ajankäytön kirjanpitoa otokseen poimituilla maatiloilla. Koska havaintoaika on kokonainen vuosi, pyritään otosta usein vaihtamalla välttämään kieltäytymisestä ja väsymisestä johtuva kato, joka on otannan sovellutuksen pahimpia esteitä tämänluonteisissa tutkimuksissa.

Kirjanpitomenetelmään verrattuna ajankäytön haastattelu on nopeampi ja useimmissa tapauksissa halvempi mittaamenetelmä. Vaikeudet sen soveltamisessa otokseen ovat huomattavasti pienemmät. Kun siinä tiedot perustuvat muistiin, on sen mittaustarkkuudessa ja jaottelun yksityiskohtaisuudessa tai vaihtoehtoisesti havaintoajan pituudessa tinnittävä vaatimuksista. Sen sijaan havaintoajan pituus sellaisenaan vaikuttaa kustannuksiin verrattain vähän. Kun haastatteliija saapuu koehenkilön luo, hän samalla tiedustelee lyhyemmän tai pitemmän ajanjakson käyttöä.

Yleensä ajankäytön haastattelututkimuksissa havaintoaika on kuitenkin ollut varsin lyhyt, tavallisimmin viikon pituinen, kuten USA:n työvoiman tutkimuksissa. Tanskan työvoiman tutkimuksessa v. 1951 se oli vain päivä (Arbejdsstyrken ... 1952). Sen sijaan näytesyksiköiden määrä on tavallisesti verrattain suuri, USA:ssa noin 25 000 ja Tanskassa noin 6 000 ruokakuntaa, mikä merkinnee suunnilleen kaksinkertaista määrää koehenkilöitä.

Viime aikoina on sekä kulutuksen haastattelututkimuksissa (esim. Fyysillisten ... 1951) että niihin monessa suhteessa verrattavissa ajankäytön haastattelututkimuksissa ruvettu käyttämään pitkiä, jopa vuoden havaintoaikoja. Jo Suomen maatalouslaskennassa v. 1941 kokeiltiin vuoden havaintoaikaa maatalojen työpanoksen mittaamisessa (PIHKALA 1947 ja Yleinen maatalouslaskenta ... 1945), ja samaa menetelmää käytettiin myös vuoden 1950 maatalouslaskennassa (PIHKALA 1952 ja Yleinen maatalouslaskenta ... 1954). Edellä mainittu Norrbottenin työvoimatutkimus v. 1943/44 oli yksityisten työllisten ajankäytön haastattelu (WIESLANDER 1948, ss. 5—9). Siinäkin havaintoaika oli vuosi, mutta mainituista maatalouslaskennoista poiketen ajankäyttö mitattiin kausittain jaotellen ja niin, että merkittiin muukin ajankäyttö kuin työpanos. Vuoden kulutuksen ja vuoden ajankäytön haastattelututkimuksissa näytesyksiköiden lukumäärä on ollut muutamasta sadasta kymmeneen tuhansiin (Norrbottenissa n. 30 000 koehenkilöä). Vuoden ajankäytön haastattelututkimuksissa mittayksikkö on ollut päivä tai arkipäivä, ja toimintojen jaottelussa on pyritty vain tärkeimpiin pääryhmiin.

Jonkin perusjoukon koko vuoden ajankäyttö on haastatteleamalla mitattavissa kausivaihteluineen sekä vuoden että sitä lyhyempääkin havaintoaikaa käyttäen. Viimeksi mainitussa tapauksessa haastatteluja samasta tai vaihtuvasta otoksesta on suoritettava yhtä suuri määrä, kuin vuodessa on havaintoaikoja. Sitä paitsi on ajateltavissa, että jokin osa havaintoajoista jätetään kokeellisesti tutkimatta ja tulokset niiden

osalta arvioidaan interpoloimalla. Vuoden havaintoaika sekä yksi haastattelu ja yksi otos on epäilemättä halvin mahdollinen menetelmä.

Vielä on eräs seikka, joka nimenomaan metsätyövoiman tutkimuksessa puoltaa vuoden todellista havaintoaikaa identtisiä koehenkilöitä käyttäen ja siten lähinnä vuoden ajankäytön haastattelua (tai kirjanpitoa). Perusjoukon metsätyöpanos, sen jakautuminen alaryhmiin (hakkuu-, ajo- ym. työpanos), maataloustyöpanos, teollisuustyöpanos jne. saadaan kausivaihteluineen arvioiduksi millä tahansa edellä mainituista menetelmistä. Samoin on työttömyysajan laita. Metsätyövoiman suuruus sekä sen ja sen alaryhmien jakautuminen työpanoksen suuruuden mukaan jonakin vuonna voidaan kuitenkin saada selville vain, jos tunnetaan samojen koehenkilöiden koko vuoden metsätyöpanos alaryhmineen. Viimeksi mainittu tieto on myös välttämätön, jotta voitaisiin tutkia ammattimaisten metsätyöntekijöiden olosuhteita ja niihin vaikuttavia tekijöitä, kuten asunto-oloja, tuloja ja maanomistusoloja, jos mainittu työvoima määritellään vuoden metsätyöpanoksen suuruuden perusteella.

Edellä ei ole kosketeltu postikyselymenetelmää ajankäytön mittauksessa. Tämä USA:ssa taloudellisissa tutkimuksissa yhä laajemman käytön saanut tiedustelumenetelmä on väestön otoksissa toistaiseksi Pohjoismaissa verrattain tuntematon. Sen heikkous on erittäin suuri vastausten kato (The Agricultural ... 1949, ss. 31—34), joka kuitenkin uusin menetelmin on yhä paremmin hallittavissa (esim. HANSEN ja HURWITZ 1946). Sen erittäin merkittävä etu on aineiston halvempuus haastattelumenetelmäänkin verrattuna. Tekijän tiedossa ei ole tapausta, jossa postitiedustelua olisi aiemmin käytetty ajankäytön otantatutkimuksessa. Sen onnistuminen vaatii, että väestö ensin totutetaan tähän kyselymenetelmään.

2. Esitutkimukset

21. Esitutkimusten järjestely

211. I esitutkimus. Haastattelututkimukset Ranuan ja Kontiolahden kunnissa v. 1949

Ennen koko valtakuntaa peittävän metsätyövoiman suuruutta, rakennetta, ajankäyttöä ja sosiaalisia oloja selvittävän tutkimuksen aloittamista päätettiin suorittaa eräitä esitutkimuksia. Niiden ensisijaisena tarkoituksena oli etsiä mainittuun tutkimukseen parhaiten soveltuva mittaus- ja otoksen poimintamenetelmä.

Edellä esitetyssä tarkastelussa todettiin, että metsätyövoiman tutkimus oli parhaiten suoritettavissa maalaiskuntien työkykyisessä iässä olevasta miespuolisesta väestöstä poimitulla otoksella. Edelleen todettiin, että koehenkilöiden ajankäytön mittauksessa saattoi tulla kysymykseen joko haastattelu- tai kirjanpitomenetelmä.

Päämääräksi asetettiin, että valtakunnan maaseudun miestyövoiman työpanos ja sen rakenne — ennen kaikkea metsätyöpanoksen suuruus ja rakenne — voitaisiin arvioida vuoden ajalta. Samalla pyrittiin selvittämään po. työvoiman työttömyysajan, myös loppoajan määrä sekä metsä- ja uittotyövoiman suuruus ja rakenne.

Koska ajankäytön mittausmenetelmistä Norrbottenissa sovellettu vuoden ajankäytön muistinvarainen tiedustelu oli tiettävästi nopein ja taloudellisin, päätettiin kokeilla sitä. WIESLANDER (1948, s. 9) tosin mainitsee, että tämän menetelmän tulosten tarkkuutta oli ennen sen sovellutusta Ruotsissa tutkittu »useilla eri tavoin toimeenpannuilla pistokokeilla», jolloin todettiin, että haastatteluissa »työllisyydestä annetut tiedot tarkkuudeltaan täyttävät suuret vaatimukset».¹ Hän selostaa kuitenkin vain yhden näistä vertailevista luotettavuustutkimuksista mainiten, että 50 miehen ajankäyttötiedusteluissa ilmoittamien uittotyöpanosten summaa verrattiin uittotyöpanosten antamiin tietoihin asianomaisten miesten yhtei-

¹ Suomennos tekijän.

sestä työpanoksesta. Uittoyhdistyksen ilmoittama työpanos ylitti 4,5 %:lla koehenkilöiden ilmoituksista lasketun vastaavan luvun. Muista Ruotsissa suoritetuista esitutkimuksista, joihin WIESLANDER viittaa, tekijä ei ole onnistunut saamaan tietoja. Näin ollen katsottiin aiheelliseksi tarkistaa myös Suomen olosuhteissa tämän mittausten menetelmän luotettavuutta, tehdä tarpeelliseksi havaittavat korjaukset haastattelulomakkeeseen ja haastattelun suoritustapaan sekä samalla valita suunniteltua otantatutkimusta varten sopiva poimintamenetelmä.

Tutkimusalueiksi valittiin Ranuan kunta Lapin läänissä ja Kontiolahden kunta Kuopion läänissä. Tutkittavien kuntien valinta oli harkinnanvarainen. Ranua valittiin edustamaan sellaista teollistumatonta kuntaa, jossa viljelmät ovat pieniä ja jossa valtaosa miespuolista, työkykyistä väestöä osallistuu metsätöihin. Kontiolahden kunta, jossa sijaitsee suuri Ristisaaren erottelulaitos ja jonka väestö työskentelee osittain myös Joensuun kaupungissa, edusti Ranuaa enemmän kaupunkimaistunutta yhteiskuntaa. Silti Kontiolahdellakin metsä- ja uittotöillä on väestön toimeentulolle huomattava merkitys. Molemmissa kunnissa arvioitiin ajankäytön haastattelun olevan yleensä vaikeinta luokkaa. Niissä suoritettiin samalla myös maatalouslaskennan esitutkimus, mikä osaltaan vaikutti kuntien valintaan.

Koekunnissa osoittautui 1. 1. 1949 suoritettujen henkikirjoituksen perusteella laadittu henkilöluettelo, ns. henkikirja, parhaaksi kehikoksi (perusluetteloksi). Sen perusteella poimittiin arpomalla kylittäin eri suuruiset otokset kaikista vv. 1883—1931 syntyneistä, henkikirjoihin merkityistä miehistä. Kontiolahdella näyte peitti vain Kunnasniemen ja Jakokosken kylät sekä Lehmon kansakoulupiirin (jota seuraavassa nimitetään Lehmon kyläksi). Ranuan näyte peitti koko kunnan. Ranuan kunnan Ylimaan kylässä sekä Kontiolahden Kunnasniemen kylässä otettiin näytteeseen kaikki perusjoukon miehet. Muissa kylissä arvottiin koehenkilöt osittamalla perusjoukko kuuteen ikäluokkaryhmään, jotka vielä jaettiin kahtia, naimisissa oleviin ja muihin. Kustakin ositteesta arvottiin otantaosuuden edellyttämä määrä, Kontiolahden Jakokosken kylässä lisäksi kullekin koehenkilölle yksi varamies. Perusjoukon suuruus, näyteyksiköiden määrä ja otantaosuudet näkyvät taulukosta 1.

Näytteeseen poimituista koehenkilöistä täytettiin henkilökohtaisin haastatteluin joukko lomakkeita, joista tärkein oli liitteissä jäljennetty ajankäyttölomake. Muihin lomakkeisiin otettiin tiedot koehenkilön perhesuhteista, iästä, asunnosta yms. sekä koehenkilön viljelemästä maatilasta, milloin sellainen oli olemassa. Kirjoittajan lisäksi oli 5 haastattelijaa. Haastattelut suoritettiin heinä- ja elokuussa 1949.

Taulukko 1. Perusjoukon suuruus, teoreettinen otantaosuus sekä näytteen suuruus. I esitutkimus.

Table 1. Size of population, theoretical sampling ratio and size of sample. Pilot Survey I.

Kunta — Kylä <i>Commune — Village</i>	Perusjoukossa miehiä <i>Males in population</i>	Teoreettinen otantaosuus <i>Theoretical sampling ratio</i>	Näytteessä miehiä <i>Males in sample</i>				Yhteensä <i>Total</i>
			Hyväksyty lomake <i>Form approved</i>	Hylätty lomake <i>Form rejected</i>	Työkyvyttömiä <i>Disabled</i>	Tietoja ei saatu <i>Non-response</i>	
Ranua — Ylimaa ..	410	1	355	11	18	26	410
Ranua ..	496	0.2	92	2	5	0	99
Simojärvi	280	0.2	51	2	3	0	56
Kemihaara	194	0.2	38	0	1	0	39
Kontiolahti — Kunnasniemi	81	1	77	0	2	2	81
Jakokoski	162	0.2	31	0	1	0	32
Lehmo ..	320	0.1	31	0	1	0	32
Yhteensä — Total	1 943	.	675	15	31	28	749

Kuten taulukosta 1 ilmenee, saatiin tiedot ajankäytöstä kerätyiksi yhteensä 721 koehenkilöltä. Loput 28 koehenkilöä olivat kuolleet, muuttaneet pois paikkakunnalta, tai heitä ei muista syistä voitu haastatella. Kieltäytymisen vuoksi tiedot jäivät saamatta vain yhdessä tapauksessa. Täytetyistä lomakkeista oli 15 puutteellisuuksien takia hylättävä ja lisäksi jätettiin 31 lomaketta käsittelemättä sen tähden, että koehenkilöt olivat ilmoittaneet olleensa vähintään puolet tutkimusvuodesta (1. 7. 1948—30. 6. 1949) työkyvyttöminä. Saadusta aineistosta käsiteltiin siis 675 ajankäyttölomaketta.

Käsittely suoritettiin kylittäin ja niissä yhteiskuntaluokittain laske- malla kussakin luokassa lomakkeiden vastaavat päivien luvut yhteen (käsin). Yhteiskuntaluokitus suoritettiin niin, että ensin erotettiin koko tutkimusvuoden ajan vakinaisessa palkatussa työssä tai yksinäisyrittäjinä (ei kuitenkaan maataloudessa) toimineet omaksi ryhmäkseen. Loput jaettiin tilattomiin sekä viljelijäperheen jäseniin, viimeksi mainitut edelleen vanhojen tilojen ja uudistilojen viljelijäperheen jäseniin. Lopuksi jaettiin vielä vanhojen tilojen viljelijäperheen jäsenet viljelmän peltopinta-alan perusteella 8 luokkaan. Yhteiskuntaluokkia tuli täten kaikkiaan 11, Kontiolahdella ne oli kuitenkin yhdistettävä 5 luokaksi. Luokituksen tarkoi-

tuksena oli pyrkiä yhteiskuntaryhmiin, joiden ajankäyttö rakenteeltaan mahdollisimman paljon eroaisi toisistaan. Luokitus tuli voida suorittaa ajankäytön rakennetta ennakolta tuntematta. Tässä esitutkimuksessa ryhmitys suoritettiin koehenkilöiltä saatujen muiden kuin ajankäyttö-tietojen perusteella. I esitutkimuksen otoksen jakautuminen yhteiskuntaryhmiin näkyy taulukosta 2.

Tämän jälkeen tulokset laskettiin yhteen molemmissa kunnissa erikseen kylittäin ilman yhteiskuntaluokitusta sekä koko kunnalle yhteiskuntaryhmiin luokitellen ja ilman sitä (suurennustekijöillä punniten). Ranuan Ylimaassa kylän aineistosta poimittiin lisäksi arpomalla yksi 20 %:n ja kuusi 10 %:n ikäluokittain ositettua rinnakkaista alanäytettä, joille kullekin laskettiin arviot vielä erikseen (luokittelematta). Lopuksi otettiin koko aineistosta kylittäin ositettu alanäyte, jonka teoreettinen otantaosuus oli 5 % perusjoukosta, ja siitä laskettiin tulokset luokittelematta.

Taulukko. 2. Miestyövoiman jakautuminen yhteiskuntaryhmiin.
I esitutkimus.

Table 2. Distribution of male labour force into social groups. Pilot Survey I.

Yhteiskuntaryhmä Social group *	Ranuan kylät Ranua villages				Koko Total Ranua	Kontiolahden kylät Kontiolahti villages			Koko Total Kon- tiolahti
	Y	R	S	K		K	J	L	
	%								
1. Vakinaisessa palk- kатыösuhteessa olevat ja itsenäis- yrittäjät (paitsi maanviljelijät) ..	5	16	4	5	8	9	6	32	21
2. Tilattomat ja 0— 0.24 ha peltoa vil- jelevät	9	18	31	18	19	12	6	16	13
3. Uudisviljelijät ..	12	7	2	7	7	10	0	16	11
4. 0.25—1.99 peltoa viljelevät	7	11	6	30	12	27	42	26	31
5. 2 ha tai enemmän peltoa viljelevät	67	48	57	40	54	42	46	10	24
Yhteensä — Total	100	100	100	100	100	100	100	100	100

- * 1. Permanently employed workers and independent entrepreneurs (excl. farmers).
2. Men without land and farmers with 0—0.25 ha under plough.
3. Farmers of newly reclaimed holdings.
4. Farmers with 0.25—1.25 ha under plough.
5. Farmers with over 2 ha under plough.

I esitutkimuksen aineiston käsittelyllä pyrittiin tutkimaan työpanosten ja työttömyysajan hajontaan vaikuttavia tekijöitä otannan osittamismahdollisuuksien ja otoksen suuruuden selvittämiseksi. Toisaalta haluttiin, vertaamalla saatuja tuloksia muihin käytettävissä oleviin tietoihin ajankäytön rakenteesta ja siihen vaikuttavista tekijöistä, todeta haastatte- luissa annettujen vastausten mahdollinen harhaisuus.

Vastausten luotettavuutta yritettiin selvittää myös siten, että haastat- telijat Ranualla merkitsivät harkintansa mukaisen ajankäyttötietojen luotettavuusasteen. Merkittävämpi oli yritys verrata I esitutkimuksessa saatuja tietoja koehenkilöiden ansiotyöpanoksesta vastaaviin tietoihin asianomaisten verokirjoissa ja palkanpidätystodistuksissa. Tätä vertaile- vaa työpäivätutkimusta varten käytettiin koko I esitutkimuksen käsiteltä- vää otosta vakinaisessa työsuhteessa olleita koehenkilöitä lukuun ottamatta. Poikkeuksen muodosti Ylimaassa kylä, jossa tyydyttiin 20 %:n otokseen. Näiden koehenkilöiden veroilmoitukset vuosilta 1948 ja 1949 tarkastet- tiin, ja verokirjoihin tai palkanpidätystodistuksiin merkityt työskentely- ajat tutkimusvuoden ajalta jäljennettiin.

212. II esitutkimus. Ranuan kirjanpitolitkimus v. 1950

Ajankäytön kirjanpidon mahdollisuuksien, ajankäytön rakenteen vai- htelun ja haastatteluvastausten virheellisyyksien tutkimiseksi poimittiin 10 %:n otos Ranuan henkikirjaan (1. 1. 1949) merkityistä vv. 1885—1933 syntyneistä miehistä. Otos poimittiin tälläkin kerralla I esitutkimuksen näytteestä kylittäin ja ikäluokittain. Siitä poistettiin ennen v. 1885 syn- tyneet ja siihen lisättiin suoraan henkikirjasta arvotut 10 % vv. 1932—33 syntyneistä. Näin saatiin 133 työkykyistä miestä käsittävä näyte.

II esitutkimuksen näytteeseen kuuluville koehenkilöille jaettiin vuo- den 1950 aikana kunkin kuukauden alussa kirjanpitokalenteri, johon hei- dän tuli päivittäin rastein päivän tarkkuudella merkitä ajankäyttönsä. Halutessaan he saivat käyttää myös puolen päivän tarkkuutta (vino- viiva). Osa eräästä aineiston kirjanpitokalenterista on jäljennettynä liit- teissä.

Kirjanpitolitkimusta varten oli järjestettävä tutkimusvuoden ajaksi pysyvä paikallinen keräilijäverkosto. Keräilijöinä toimivat etupäässä Työ- keskus Rovalan opintokerholaiset siten, että kunkin tehtävänä oli huo- lehtia 2—5 koehenkilöstä. Keräilijöiden valvojina toimi kymmenkunta ohjaajaa (useimmat kansakoulunopettajia). Keräilyverkoston tehtävänä

oli joka kuukauden alussa jakaa Helsingistä lähetetyt kalenterit koehenkilöille, kerätä heiltä edellisen kuukauden kalenterit sekä samalla tarkistaa, että viimeksi mainitut oli oikein täytetty. Mikäli ilmeni puutteellisuuksia, kalenteri tuli täydentää haastattelemalla. Kun mainitusta keräilyhenkilökunnasta osa pyrki varsinkin aluksi keskeyttämään toimintansa ja koehenkilötkin siirtyivät paljon paikasta toiseen, oli tutkimusvuoden aikana useaan kertaan osa kalentereista kerättävä tarkoitusta varten palkatun kiertävän kerääjän toimesta. Vuoden 1950 loppuun mennessä 21 koehenkilöä siirtyi vähitellen suoraan kirjeysteeseen Helsingin yliopiston metsätaloustieteellisen laitoksen kanssa, joten aineiston keräys heidän osaltaan muuttui lähinnä postitiedustelumenetelmää muistuttavaksi. Tämä keräysmenetelmän muutos oli joissakin tapauksissa välttämätön sen tähden, että koehenkilö matkusti toiseen kuntaan, jopa naapurimaahan. Useat koehenkilöt pitivät myös suoraa yhteyttä yliopistoon miellyttävimpänä.

Vaikkei koehenkilöille eikä keräyshenkilökunnallekaan (kiertäviä keräjiä lukuun ottamatta) voitu maksaa mainittavaa korvausta työstä, saatiin koko vuoden aineisto keräytyä kaikilta koehenkilöiltä.

Saatu aineisto käsiteltiin laskemalla päivittäin 133 arkipäivän suhteellinen jakautuminen eri ajankäyttöryhmiin. Samalla tavalla laskettiin tulokset myös kuukausittain ja koko tutkimusvuodelta.

213. III esitutkimus. Ajankäytön haastattelulomakkeiden kokeilu v. 1950

Kun neuvotteluja päätutkimuksen toimeenpanemisesta käytiin keväällä 1950, suoritettiin ajankäytön haastattelulomakkeen parhaimman muodon selvittämiseksi tutkimus erityyppisillä lomakkeilla Helsingin ja Tuusulan maalaiskunnissa. Tämä koe oli kiireellisyytensä takia supistettava hyvin pieneen aineistoon. Kaikkiaan täytettiin 48 lomaketta 28 koehenkilöstä. Koehenkilöt valittiin mielivaltaisesti. Heistä oli 19 maatalojen vakinaisia työntekijöitä, loput olivat maatalousharjoittelijoita ja viljelmän haltijoita.

Haastattelijoina toimi maatalous- ja metsätieteiden kandidaatti (yksin) sekä 2 maatalousharjoittelijaa yhdessä. Ensiksi mainittu ei saanut neuvotella tutkimuksen aikana (3 päivää) siitä muiden haastattelijoiden kanssa. Lomaketyyppejä oli 4. Niistä kahdessa vuosi oli jaettu kalenterikuukausiin. Toisessa tuli ajankäyttö merkitä päivän, toisessa 1/4 kuukauden tarkkuudella. Muissa lomakkeissa oli vuosi jaettu viiteen kauden. Toisessa

tuli kaikki työpanosryhmät merkitä kausittain, toisessa maataloustyöt oli merkittävä kausiin jakamatta, koko vuodelta. Haastattelijoiden tuli täyttää koehenkilöistä eri lomakkeita edeltä käsin määrätyn järjestelmän mukaan siten, että kutakin lomaketyyppeä tuli täytetyksi suunnilleen sama määrä, 12 kpl. Aineiston kerättyään haastattelijat kirjoittivat lausunnon kunkin lomaketyypin täyttämisen vaikeudesta ja tietojen luotettavuusasteesta harkintansa mukaan.

22. Esitutkimusten tulokset

221. Poimintamenetelmä

Esitutkimuksissa I ja II ei kunnista ja kuntiin perusluetteloinnin ajankohdan jälkeen muuttaneita koehenkilöitä kustannusten säästämiseksi yritettykään tutkia. Sen sijaan tilapäisesti tavoittamattomista koehenkilöistä I esitutkimuksessa pyydettiin tarvittavat tiedot asianomaisen perheenjäseniltä tai lähinaapureilta, milloin se oli mahdollista. Jakokosken kylässä käytettiin varamiesjärjestelmää. Tilastoteoreettisesti viimeksi mainittua tapaa ei voida pitää suositeltavana. Niissä kylissä, joissa I esitutkimuksessa tyydyttiin näytteeseen, saatiin (tosin neljässä tapauksessa liian puutteelliset) tiedot kaikista otokseen poimituista koehenkilöistä. Täydellisesti tilastoitaviksi tarkoitetuilla alueilla oli puuttuvia tietoja sen sijaan huomattava määrä. Tämä johtui pääasiassa siitä, että kustannussyistä ei täydellisesti tutkittavilla alueilla voitu ryhtyä etsimään hyvin vaikeasti tavoitettavia koehenkilöitä. Voitiin myös todeta, että pieni joukko otoksen koehenkilöitä oli helpompi 100 %:sti tavoittaa joltakin alueelta kuin alueen koko määrätynikäinen miesväestö. Useimmat arvotuista koehenkilöistä olivat kiinnostuneet tietämään, kenelle tutuistakin »arpa on langennut» ja halukkaasti neuvoivat haastattelijan matkasuunnitelmaan muutoksia. Täydellinen luettelo sen sijaan ei herättänyt läheskään samanasteista mielenkiintoa, joten turhia matkoja tuli suhteellisesti enemmän.

Näissä esitutkimuksissa poimittiin otos kylittäin, ikäluokittain ja siviilisäädyyttäin ositetusta perusjoukosta. Tutkittaessa I esitutkimuksessa koehenkilöiden työpanoslukujen hajontaa kylien sisäisessä aineistossa ja kylien keskiarvojen hajontaa kuntien keskiarvoista voitiin todeta, että kylien sisäinen hajonta oli paljon suurempi kuin niiden välinen hajonta. Molempien kuntienkin sisäinen hajonta oli yleensä niiden välistä hajontaa suurempi. Sama koski ikäluokkien sisäistä ja välistä hajontaa.

Näin ollen ei enempää alueittainen kuin ikäluokittainenkaan osittaminen mainittavasti lisännyt koko kunnalle tai molemmille kunnille yhteisesti laskettujen estimaattien varmuutta (HANSEN, HURWITZ ja MADOW 1953, s. 201).

Kun tulokset laskettiin I esitutkimuksessa myös yhteiskuntaryhmitäin, voitiin todeta, ettei niidenkään sisäinen hajonta ollut keskimäärin ainakaan huomattavasti pienempi kuin ryhmien välinen hajonta. Näin ollen ei perusjoukon osittaminen tälläkään tavalla olisi tutkituissa kunnissa tuottanut olennaista hyötyä. Käytettyä yhteiskuntaryhmitystä varten tarvittavat tiedot ovat erittäin vaikeasti hankittavissa ilman erikois-tutkimuksia. Henkikirjoihin merkittyjen ammattinimitysten perusteella sen sijaan olisi todennäköisesti mahdollista jakaa miehet pariin tutkitulta ajankäyttörakenteeltaan olennaisesti poikkeavaan ryhmään: henkisen alan ja teollisuuden työlliset yhtenä ryhmänä sekä muut toisena ryhmänä. Ranualla ja Kontiolahdella ensiksi mainittuun ryhmään olisi kuitenkin tullut aivan liian vähän koehenkilöitä, jotta tällainen ositus olisi olennaisesti lisännyt koko koealueelle tai kunnittain laskettujen estimaattien varmuutta. Kontiolahden vakinaiset työntekijät olivat pääasiassa uittotyömiehiä.

Esitutkimukset osoittivat, että Suomessa on väestöstä otantatutkimuksia varten käyttökelpoinen luettelo, henkikirjat, päinvastoin kuin esim. USA:ssa (HANSEN ym. 1953, s. 97) ja Englannissa (YATES 1949, s. 63). Näin ollen Suomessa voidaan poimia otos väestöstä vaivattomammin kuin viimeksi mainituissa ja ehkä useissa muissakin maissa. Suorittamalla yksiasteinen poiminta suoraan henkikirjoista saavutetaan vaadittu estimaattien varmuus harvinaisia poikkeuksia lukuun ottamatta pienemmällä otoksella kuin moniasteisessa (ryväs-)poiminnassa, jollaista muissa maissa väestön tutkimuksissa usein on käytettävä. Jos aineiston keräyskustannukset ovat lähimain suoraan verrannolliset otoksen suuruuteen ja jos yksiasteisenkin poiminnan mittausharhat ovat kyllin pienet, Suomessa näyttää olevan edullista tyytyä yksiasteiseen poimintaan väestön tutkimuksissa (vrt. HANSEN ym. 1953, s. 243).

Määrätty estimaatin varmuusaste ja -alue saavutetaan pienimmällä mahdollisella yksiasteisella otoksella, jos poiminta on optimaalisesti ositettu. Koekunnissa ei näyttänyt löytyvän mahdollisuuksia osittamalla lisätä työpanoslukujen keskiarvojen (tai työpanoslukujen summien) keskimääräistä varmuutta merkittävästi, jollei tuloksia haluttu laskea ositteista erikseen. Viimeksi mainitussakin tapauksessa osittamisen hyöty on kyseenalainen, jos otos on niin suuri, että tutkittaviin näytteen ala-

ryhmiin osittamattakin suurella todennäköisyydellä saadaan riittävä määrä koehenkilöitä. Sellaisten kuntaryhmien miesten ajankäyttöä tutkittaessa, joissa henkisen alan ja teollisuuden työllisten osuus on huomattavan suuri, voinee perusjoukon osittaminen henkikirjaan merkityn ammattinimikkeen perusteella olla suositeltavaa.

222. Otoksen suuruus

Päätutkimuksen otoksen suuruuden määrasivät ennen kaikkea käytettävissä olleet varat. Tämä päätös oli myöskin tehtävä niin kiireellisesti, että I ja II esitutkimuksesta ehdittiin suorittaa vain suuripiirteisiä määrättyä varmuutta varten vaadittavan otoksen suuruuden arviointeja. Selostettavat esitutkimusten hajonta-analyysit suoritettiin pääasiassa mainitun päätöksen jälkeen. Niillä on kuitenkin merkitystä päätutkimuksen varmuuslaskelmille.

Seuraavassa esityksessä käytetään alla lueteltuja näytteen tunnuslukujen merkkejä. Perusjoukon vastaavista tunnusluvuista käytetään tarvittaessa vastaavia suuria kirjaimia. Merkkijärjestelmässä noudatetaan pääasiassa HANSENIN ym. (1953) esitystä. Tärkeimmät tässä tutkimuksessa käytetyt merkit on myös selitetty ss. 137—141. Liitteisiin jäljennetyn I esitutkimuksen ajankäyttölomakkeen tarkastelu auttaneen lukijaa merkkijärjestelmän omaksumisessa.

- n näytteen suuruus eli sen koehenkilöiden ja samalla ajankäyttölomakkeiden lukumäärä.
- m näytteen määrättyä yhteisellä ominaisuudella varustettu alaryhmä, joka voi saada maksimi-arvon n .
- x_{ij} näytteen alaryhmän i :n lomakkeen j :n ajankäyttöryhmään (esim. metsätyöt, hakkuu) merkittyjen päivien lukumäärä. Kulloinkin mainitaan erikseen, onko kysymys havaintovuoden määrätystä osakaudesta (esim. käsi VI = 15. 5. — 30. 6. 1949) vai koko havaintokaudesta (-vuodesta). Paitsi alkuperäisistä ajankäyttöryhmistä, voidaan x_{ij} laskea myös niitä yhdistelemällä, esim. »yhteensä päiviä viljelmän tai tontin ulkopuolella» on eräs tällainen yhdistetty ryhmä.
- z_i eräs yhdistelemällä laskettu x_{ij} -luku, nimittäin i :n lomakkeen kaikkien määrätyn kauden alkuperäisten ja itsenäisten x_{ij} -lukujen summa.
- $a = z_i$, kun z_i on vakio.

$p_{ij} = \frac{x_{ij}}{z_i}$. Mikäli tämän luvun jälkeen seuraa sadanneksen merkki %, on luku kerrottu 100:lla. Sama koskee kaikkia suhdelukuja.

$$q_{ij} = 1 - p_{ij}.$$

$x_j = \sum_i^m x_{ij}$. Näytteen alaryhmän ajankäyttöryhmän j kaikkien x_i -lukujen summa.

$\bar{x}_j = \frac{x_j}{m}$. Näytteen alaryhmän ajankäyttöryhmän j kaikkien x_i -lukujen keskiarvo.

$z = \sum_i^m z_i$. Kun $z_i = a$, on $z = ma$.

$\bar{p}_j = \frac{x_j}{z}$. Jos $z_i = a$, on \bar{p} myös $= \frac{\bar{x}_j}{a} = \frac{\sum_i^m p_{ij}}{m}$, siis kaikkien lomakkeiden p_{ij} -lukujen keskiarvo.

$$\bar{q}_j = 1 - \bar{p}_j.$$

$F = \frac{n}{N}$. Otantaosuus.

$K = \frac{N}{n}$. Suurennustekijä.

$x'_j = Kx_j$. Näytteestä laskettu X_j -luvun estimaatti.

$z' = Kz$. Näytteestä arvioitu Z . Kun $z_i = a$, on $z' = Kma$.

$s_{x_{ij}}^2$ x_{ij} -lukujen (absoluuttinen) varianssi eli keskihajonnan neliö.

$s_{\bar{x}_j}^2$ näytteestä arvioitu (perusjoukon kaikkien mahdollisten samanlaatuisten ja -suuruisten rinnakkaisten näytteiden) \bar{x}_j -estimaattien varianssi eli \bar{x}_j -estimaatin (absoluuttisen) keskivirheen neliö.

$v_{x_{ij}}^2 = \frac{s_{x_{ij}}^2}{\bar{x}_j^2}$. Lukujen x_{ij} näytteestä laskettu suhteellinen varianssi eli variaatiokertoimen neliö.

$v_{\bar{x}_j}^2 = \frac{s_{\bar{x}_j}^2}{\bar{x}_j^2}$. Luvun \bar{x}_j näytteestä laskettu suhteellinen varianssi eli suhteellisen keskivirheen neliö.

$s_{p_{ij}}^2$, $s_{\bar{p}_j}^2$, $v_{p_{ij}}^2$ ja $v_{\bar{p}_j}^2$ ovat vastaavat p_{ij} -luvuista lasketut estimaatit.

$s_{x_j}^2$, $s_{x'_j}^2$, $v_{x_j}^2$ ja $v_{x'_j}^2$ ovat vastaavat x_j - ja x'_j -estimaateille näytteestä lasketut varianssit.

Edellisestä seuraa, että kun $m = n$, on $v_{x_j}^2 = v_{\bar{p}_j}^2$ ja riippumatta siitä, onko $m = n$, on $v_{\bar{p}_j}^2 = v_{x_j}^2$ kun $z_i = a$.

Seuraavassa käytetään varianssin laskennassa likiarvokaavoja, joissa äärellisen perusjoukon kerroin $(1-F)$ on jätetty huomioon ottamatta ja lisäksi edellytetään, että (YATES 1949, s. 185 ja HANSEN ym. 1953, s. 125)

$$s_{x_{ij}}^2 = \frac{\sum_i^m (x_{ij} - \bar{x}_j)^2}{m} = \frac{\sum_i^m x_{ij}^2}{m} - \bar{x}_j^2.$$

Silloin saadaan

$$v_{\bar{x}_j}^2 = \frac{\sum_i^{m_1} x_{ij}^2}{\left(\sum_i^{m_1} x_{ij}\right)^2} \frac{m_1}{m_2} - \frac{1}{m_2}, \quad (1)$$

jossa m_1 on näyteyksiköiden lukumäärä siinä (ala)otoksen määrättyllä ominaisuudella varustetussa ryhmässä, josta laskutyöt suoritetaan ja m_2 on näyteyksiköiden lukumäärä sen otoksen vastaavassa alaryhmässä, jolle $v_{\bar{x}_j}^2$ lasketaan.

Kaavasta 1 saadaan

$$m_2 = \frac{1}{v_{\bar{x}_j}^2} \left[\frac{\sum_i^{m_1} x_{ij}^2}{\left(\sum_i^{m_1} x_{ij}\right)^2} m_1 - 1 \right] \quad (2)$$

Esitutkimuksessa I oli $z_i = a$, joten kaavoissa 1 ja 2/s. 37 tällöin voidaan panna $v_{\bar{x}_j}^2$ -merkin paikalle $v_{\bar{p}_j}^2$. Mainitun tutkimuksen varianssilaskelmat perustuvat koko näytteeseen, joten kaavoissa korvataan m n :llä ja siten $v_{\bar{x}_j}^2$ on myös $= v_{\bar{x}'_j}^2$.

Taulukkoon 3 on laskettu eri ajankäyttöryhmille n_2 edellyttäen, että $v_{\bar{p}_j}^2 = 0.01$ ja siis $v_{\bar{p}_j} = \pm 10\%$. Laskelmat on suoritettu seuraavalla kolmella tavalla:

1) I esitutkimuksen koko näytteestä otettiin 5% perusjoukosta vastaava alanäyte, jonka suuruus (n_1) oli 87 lomaketta. Näistä lomakkeista laskettiin koko havaintovuoden n_2 kaavasta 2/s. 37.

2) Samasta alanäytteestä laskettiin kaudelle VI (16. 5. — 30. 6. 1949) vastaavalla tavalla sarja n_2 .

3) Ylimään kylän »täydellisestä» aineistosta otettiin 6 rinnakkaista 10%:n otosta (35 koehenkilöä kussakin) ja laskettiin koko tutkimusvuoden arviot $v_{\bar{p}_j}^2$ seuraavasti

$$v_{\bar{p}_j}^2 = \frac{\sum_l (\bar{p}_{jl} - \bar{P}_j)^2}{6},$$

jossa \bar{p}_{jl} tarkoittaa l :nnestä alaotoksesta laskettua arviota \bar{p}_j ja \bar{P}_j »täydellisestä» aineistosta laskettua vastaavaa lukua. Tämän jälkeen laskettiin luvut n_2 (sarja 3) kaavalla

$$n_2 = \frac{n_1}{0.01} v_{\bar{p}_j}^2$$

Selvyyden vuoksi käytetään näin lasketuille n_2 -luvuille seuraavassa merkkiä $n_{(v=0.1)}$.

Kuvaan 1/s. 40 on taulukosta 3 merkitty puolilogaritmiseen koordinaatistoon saadut $n_{(v=0.1)}$ -estimaatit \bar{p}_j -estimaattien funktioina sekä piirretty näin saaduille pistesarjoille silmävaraisesti tasoituskäyrät. Tulokset osoittaa, että kaikissa tutkituissa tapauksissa $n_{(v=0.1)}$ keskimäärin pieneni \bar{p}_j -luvun suuretessa ja että se oli huomattavasti suurempi $1^{1/2}$

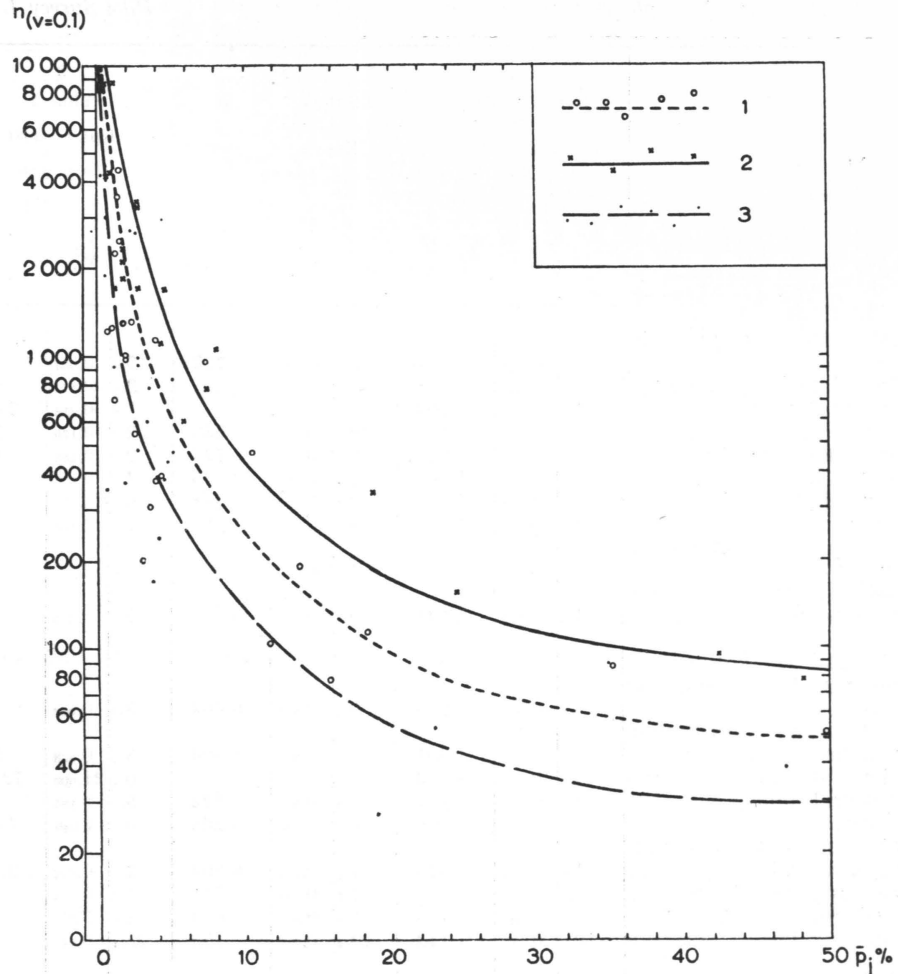
Taulukko 3. Otoksen suuruus, kun varmuusvaatimus on $v_{\bar{p}_j} = \pm 10\%$.

I esitutkimus.

Table 3. Size of sample for confidence requirement of $v_{\bar{p}_j} = \pm 10\%$. Pilot Survey I.

Ajankäyttöryhmä Activity group *	1 Ranua + Kontio- lahti Koko vuosi Total year (n = 87)			2 Ranua + Kontio- lahti Kesä — Summer 15. 5. — 16. 6. 1949 (n = 87)			3 Ylimaa Koko vuosi Total year (n = 35)		
	\bar{p}_j	$v_{\bar{p}_j}^2$	$n_{(v=0.1)}$	\bar{p}_j	$v_{\bar{p}_j}^2$	$n_{(v=0.1)}$	\bar{p}_j	$v_{\bar{p}_j}^2$	$n_{(v=0.1)}$
	1. Metsätalous, hakkuu, ansiotyö	13.8	0.022	191	2.8	0.198	1 722	9.6	0.033
2. Metsätalous, ajo, ansiotyö	3.9	0.043	374	—	—	—	3.4	0.173	606
3. Metsätalous, muu ansiotyö	1.1	0.081	704	—	—	—	2.8	0.819	2 867
4. Uitto	10.5	0.054	469	18.9	0.039	339	2.3	0.170	595
5. Maatalous, ansiotyö	2.6	0.063	548	7.5	0.089	774	2.8	0.263	921
6. Tietyö (ei työttömyystyö)	2.4	0.157	1 366	2.7	0.388	3 375	4.8	0.128	448
7. Saha- ja lastaustyö	1.8	0.151	1 314	2.7	0.378	3 288	0.6	0.875	3 062
8. Talonrakennus, ansiotyö	4.0	0.139	1 209	4.6	0.198	1 722	1.9	0.106	371
9. Muu ansiotyö (ei työttömyystyö)	7.4	0.107	931	8.1	0.122	1 061	3.6	0.226	791
10. Julkiset tehtävät, koulu, asepalvelu	1.4	0.407	3 541	0.9	0.499	4 341	2.8	0.279	977
11. Metsätalous, kotitarvepuun hakkuu ja ajo, oma työ	3.0	0.023	200	0.2	0.989	8 604	3.7	0.042	147
12. Metsätalous, myyntipuun hakkuu ja ajo, oma työ	1.9	0.115	1 001	0.4	0.989	8 604	5.1	0.228	833
13. Metsätalous, muu oma työ	0.1	0.989	8 604	—	—	—	0.1	21.000	73500
14. Talonrakennus, oma työ	4.3	0.045	392	8.8	0.060	522	5.1	0.135	473
15. Kotiteollisuus	1.6	0.288	2 506	0.5	0.490	4 263	0.7	1.320	4 620
16. Poronhoito ja liiketoiminta, oma työ	1.6	0.503	4 376	1.2	0.989	8 604	2.7	0.758	2 653
17. Metsästyys ja kalastus	0.7	0.141	1 227	1.1	0.199	1 731	1.2	0.264	924
18. Uudisraivaus, oma työ	3.5	0.035	305	5.8	0.069	600	4.6	0.107	375
19. Muu maatalous, oma työ	18.6	0.013	113	24.4	0.018	156	23.4	0.015	53
20. Työttömyystyö	0.9	0.144	1 253	—	—	—	0.6	0.550	1 925
21. Avoin työttömyys ja luppoaika	11.7	0.012	104	6.0	0.094	818	15.0	0.012	42
22. Sairaus	1.9	0.115	1 001	1.7	0.211	1 835	2.7	0.138	483
23. Vapaa-aika	1.3	0.258	2 245	1.7	0.273	2 375	0.5	0.100	350
24. Yhteensä ryhmät	100.0	.	0	100.0	.	0	100.0	.	0
25. Yhteensä ryhmät	48.9	0.006	52	48.2	0.009	78	34.6	0.025	88
26. Yhteensä ryhmät	35.2	0.010	87	42.4	0.011	95	46.6	0.011	39
27. Yhteensä ryhmät	15.9	0.009	78	9.4	0.044	382	18.8	0.008	28

* 1. Forestry, felling, paid work. — 2. Forestry, hauling, paid work. — 3. Forestry, other paid work. — 4. Floating. — 5. Farming, paid work. — 6. Road construction (excl. relief work). — 7. Sawmill and loading work. — 8. House construction, paid work. — 9. Other paid work (excl. relief work). — 10. Public activities, schools, military service. — 11. Forestry, felling and hauling of wood, for domestic use, unpaid work. — 12. Forestry, felling and hauling of wood for sale, unpaid work. — 13. Forestry, other unpaid work. — 14. House construction, unpaid work. — 15. Domestic handicraft. — 16. Reindeer raising and business activities, unpaid work. — 17. Hunting and fishing. — 18. Land clearing, unpaid work. — 19. Other unpaid work in agriculture. — 20. Relief work. — 21. Unemployment and underemployment. — 22. Illness. — 23. Leisure.



Kuva 1. Otoksen suuruus, kun varmuusvaatimus on $v_{\bar{p}_j} = 10\%$. I esitutkimus.

1) Ranuan ja Kontiolahden aineisto, koko vuosi 1948/49. 2) Sama aineisto, kesäkausi 15. 5. — 30. 6. 1949. 3) Ylimaän aineisto, koko vuosi 1948/49. Lähde: taulukko 3/s. 39. Puolilogaritminen asteikko.

Figure 1. Size of sample for confidence requirement of $v_{\bar{p}_j} = 10\%$. Pilot survey I.

1) Based on material from Ranua and Kontiolahti, total year 1948/49. 2) The same material, summer period May 15th — June 30th. 1949. 3) Based on material from Yli-maa, total year 1948/49. Source: Table 3/p. 39. Semi-logarithmic scale.

kuukauden kaudelle kuin koko vuodelle I esitutkimuksen koko alueen peittävästä otoksesta laskettuna. Koska $s_{\bar{p}_j}^2 = s_{q_j}^2$, voidaan nähdä, että tasoituskäyrät laskisivat \bar{p}_j :n suuretsa yli 50 %:n, kunnes pisteessä $\bar{p}_j = 100\%$ tulee $n_{(v=0.1)} = 0$.

Noin 1 000 koehenkilön yksinkertaisella otoksella on siis odotettavissa perusjoukosta, jonka ajankäytön homogeenisyys on sama kuin Ylimaän kylässä (hyvin homogeeninen), n. 2 % suurempien koko vuoden \bar{p}_j -estimaattien suhteellisen keskivirheen olevan korkeintaan 10 %. Perusjoukosta, jonka homogeenisyys on sama kuin koko I esitutkimuksen alueella (varsin heterogeeninen), voi samansuuruisen otoksen odottaa antavan koko vuoden \bar{p}_j -estimaatit, jotka ovat $> 4\%$ ja 1 1/2 kuukauden \bar{p}_j -estimaatit, jotka ovat $> 7\%$ enintään $\pm 10\%$:n varmuusvälein ($v_{\bar{p}_j} = \pm 10\%$). Jotta varmuusvaatimus $v_{\bar{p}_j} = \pm 5\%$ tulisi vastaavissa tapauksissa täytetyksi, tulisi otoksessa olla n. 4000 koehenkilöä, jolloin jo koko vuoden $\bar{p}_j > 1\%$ saisi suunnilleen $\pm 10\%$:n keskivirheen. Suurehkoille koko vuoden \bar{p}_j -estimaateille, tapauksessa 1), kun $\bar{p}_j > n. 20\%$ ja tapauksessa 3), kun $\bar{p}_j > n. 12\%$ sen sijaan voidaan odottaa saavutettavan tällöin varsin tyydyttävä $\pm 10\%$:n keskivirhe jo n. 100 koehenkilön näytteellä. On merkittävää, että I esitutkimuksen koko alueen näytteen perusteella työttömyysajan koko vuoden \bar{p}_j -estimaateille 11.7, % jo 100 koehenkilön näytteellä voidaan odottaa $\pm 10\%$:n keskivirhettä.

Myös II esitutkimuksen aineistosta laskettiin vertailun vuoksi $n_{(v=0.1)}$ -estimaatit helmikuun luvuille \bar{p}_j . Saatujen lukujen avulla piirretty tasoituskäyrä osoittautui seuraavan käyrää 2) kuvassa 1.

223. Mittausmenetelmä

Esitutkimusten I ja II eräänä tärkeimpänä tavoitteena oli ajankäytön mittausmenetelmien luotettavuuden tutkiminen. Kokonaisen vuoden ajankäytön haastattelussa on odotettavissa vastauksissa olevan runsaasti varsin huomattavia epätarkkuuksia. Tosin vastausten satunnaiset virheet pyrkivät aineiston suuretsa yhä enemmän tasoittamaan toisensa. Niitä voidaan nimittää vastaus- tai mittaushajonnaksi. Sen vaikutus lopputulosten varmuuteen tulee näkyviin tulosten hajonta- (keskivirhe-)laskelmissa, ja siten se korkeintaan lisää määrättyä tulosten varmuutta varten vaadittavan otoksen suuruutta.

Tulosten kannalta todellisen vaaran muodostavat sen sijaan systemaatt-

tiset mittausvirheet eli mittausharhat. Niitä ei voida itse otoksesta, enempää kuin täydellisen tutkimuksen aineistostakaan, todeta matemaattisin keinoin eikä otoksen suuruutta lisäämällä vähentää niiden vaikutusta tuloksiin (vrt. HANSEN ym. 1953, ss. 89—91).

Esitutkimuksessa II käytetty ajankäytön kirjanpitomenetelmäkään ei ole virheetön mittaustapa. Ei ollut mahdollista todeta, kuinka monet koehenkilöistä täyttivät kuukauden kirjanpitokalenterinsa vasta kuukauden päätyttyä. Ainakin niissä varsin lukuisissa tapauksissa, joissa koehenkilö ei suorittanut pyydettyjä merkintöjä ensinkään, vaan kalenteri oli haastatteleamalla täytettävä, saattoi muistivirheitä esiintyä. Virhemerkintöjä on luonnollisesti voinut tulla, vaikka kirjaa työpäivistä olisi pidetty päivittäinkin.

Esitutkimusten mahdolliset mittausharhat pyrittiin toteamaan seuraavilla tavoilla:

- 1) Vertaamalla I esitutkimuksen tuloksia verokirja-aineistoon.
- 2) I esitutkimuksen haastattelijat tekivät arviointeja vastausten luotettavuudesta ja siihen vaikuttavista tekijöistä.
- 3) Tutkimalla saatujen tulosten »järkevyyttä».

Luvussa 211 mainitussa vertailevassa työpäivätutkimuksessa saatiin tiedot Ranualla 194 ja Kontiolahdella 118 koehenkilöltä. Ranuan tutkimuksen koehenkilöt olivat haastatteluissa ilmoittaneet yhteensä 14 % suuremman ansiotyöpanoksen kuin verokirjoista ja palkanpidätystodistuksista laskettu ansiotyöpanos osoitti. Kontiolahden vastaavansuuntainen ero oli vain 1 %. Taulukosta 4 selviää Ranuan kunnan osalta tämän eron vaihtelu havaintovuoden eri kausina. Näyttää siltä, ettei se seikka, kuinka kaukaisesta ajanjaksosta haastattelussa oli kysymys, ainakaan selvästi vaikuta mainittuun suhteeseen.

Samaa asiaa Ranuan aineistossa tutkittiin myös yhteiskuntaryhmitäin. Keskimukoa suurempien (peltoa yli 5 ha) viljelmien haltijaperheen jäsenillä oli keskimäärin haastattelutiedoissa pienempi ansiotyöpanos kuin verokirjoissa ja palkanpidätystodistuksissa. Tilattomilla ja pienten tilojen haltijoilla suhde oli päinvastainen. Tämä voinee johtua mm. siitä, että tällaisilla »suurehkojen» tilojen viljelijöillä suuremmissa määrin kuin muilla on tapana esim. metsätyömaalta käydä hoitamassa tilansa asioita ja silloin muutamaksi päiväksi keskeyttää ansiotyö. Kun tällaisia keskeytyksiä useimmiten ei merkitä verokirjaan, tulee sen perusteella laskettu ansiotyöpanos todellista suuremmaksi. Samaan syyhyn viittaa myös se, että Kontiolahden aineiston perusteella todettiin niillä alueilla, joissa omien maatilatalouden töiden merkitys oli suurin (Jakokosken ja Kunnas-

Taulukko 4. Ansiotyöpanos haastattelujen sekä verokirjojen mukaan. Ranua. I esitutkimus. 194 koehenkilöä.

Table 4. Input of paid labour, according to interviews and tax books. Ranua. Pilot Survey I. 194 males in sample.

Havaintokausi Period of observation	Ansiotyöpanos, päiviä Input of paid labour, days		$\frac{A}{B}$ %
	Haastattelujen mukaan According to interviews A	Verokirjojen mukaan According to tax books B	
	1. 7.— 30. 9. 1948	3 806	
1. 10.— 15. 11. 1948	2 131	2 111	101
16. 11.— 31. 12. 1948	1 813	1 318	138
1. 1.— 31. 3. 1949	7 142	6 421	111
1. 4.— 15. 5. 1949	2 221	1 804	123
16. 5.— 30. 6. 1949	2 277	1 882	121
Yhteensä — Total 1948/49	19 390	16 981	114

niemen kylät), veroilmoituksista saatavan suurempi ansiotyöpanos kuin haastattelulomakkeista. Sen sijaan Lehmon kansakoulupiirissä, jonka koehenkilöistä valtaosa eli yksinomaan ansiotulojensa varassa (uittoyhdistyksen vakinaisia työntekijöitä), mainittu suhde oli päinvastainen, kuten taulukosta 5/s. 44 selviää. Samalla myös ilmeni, että verokirjojen ja palkanpidätystodistusten perusteella laskettu keskimääräinen työpäivän ansio oli Kunnasniemen koehenkilöillä kohtuuttoman pieni (431 mk/päivä), Jakokosken aineistossa taas paremmin kyseisen vuoden päiväpalkkatasoa vastaava (564 mk/päivä). Tämäkin vahvistaa käsitystä, että mainitunlaisissa oloissa ansiotyöaika tuli merkityksi verokirjoihin todellista pitemmäksi.

Myöskään sen seikan, että haastattelutietojen ansiotyöpanos ylittää veroilmoituksista saadun ansiotyöpanoksen, ei tarvitse osoittaa koehenkilön antaneen virheellisiä tietoja haastattelijalle. On luultavaa, että kaikkia ansiotöitä ei merkitä verokirjoihin eikä palkanpidätystodistetta aina liitetä veroilmoitukseen. Toisaalta on maaseudulla tapana tehdä palkkomiakin työpäiviä naapurin hyväksi (esim. ns. työvaihtopäivät), jotka haastattelussa tulivat merkityiksi, mutta joita veroviranomaisille ei tarvitsekaan ilmoittaa.

Näin ollen ei verokirjaa eikä palkanpidätystodistusta voida pitää

Taulukko 5. Ansiotyöpanos haastattelujen sekä verokirjojen mukaan. Kontiolahti. I esitutkimus.

Table 5. Input of paid labour, according to interviews and tax books. Kontiolahti. Pilot Survey I.

Tutkimusalue District	Näytteessä miehiä Males in sample	Ansiotyöpanos, päiviä Input of paid labour, days		A B %
		Haastattelujen mukaan According to interviews A	Verokirjojen mukaan According to tax books B	
Lehmo	45	11 505	10 732	107
Jakokoski	35	6 055	6 119	99
Kunnasniemi	38	4 824	5 305	91
Yhteensä — Total	118	22 384	22 156	101

todellista ansiotyöpanosta ilmaisevana mittapuuna, jonka avulla haastattelututkimuksen mittausharhat varmasti voidaan todeta. Selostetut vertailut osoittavat vain, ettei niissä tullut ilmi seikkoja, jotka todistaisivat koehenkilöiden antamien tietojen keskimäärin poikkeavan todellisuudesta. Pikemminkin ne osoittivat heidän varsin luottavaisesti pyrkineen antamaan mahdollisimman tarkat tiedot vieraalle tekemiensä työpäivien määristä.

Esitutkimuksessa I tuli Ranualla toimineiden haastattelijoiden merkitä ajankäyttölomakkeen liitteeseen, oliko tiedustelun tulos heidän haastattelun kulusta saamansa käsityksen mukaan »luotettava», »keskinkertaisen luotettava» vai »epäluotettava». Ylimään kylän 362 haastatellusta luokiteltiin 37 % ensiksi mainittuun, 60 % toiseen ja 3 % epäluotettavien ryhmään. Mainittakoon, että paikalliset kaksi aineiston kerääjää, jotka tunsivat useimmat koehenkilöt, pitivät tuloksia keskimäärin luotettavampina kuin kaksi vieraspaikkakuntalaista kerääjää. Kun koehenkilön vastauksista ja keskusteluista hänen kanssaan voidaan usein todeta, palautuvatko ajankäyttöä koskevat tiedot muistiin helposti vai vaivalloisesti, osoittavat em. tulokset, että saadut vastaukset olivat ainakin haastattelijoiden mielestä kohtalaisen todenmukaiset.

Esitutkimuksissa I ja III tehtiin myös havaintoja siitä, millä tavalla koehenkilön ammatti vaikuttaa muistamisen helppouteen ja samalla vastausten luotettavuuteen. Näiden havaintojen perusteella koehenkilöt luo-

kiteltiin seuraaviin »vaikeusryhmiin». Ryhmitys edellyttää, että ajankäytön jaottelu on suurin piirtein sama kuin esitutkimuksissa. On nimitäin huomattava, että ruumiillisen työn panos on paljon helpompi jakaa muistinvaraisesti ajanjaksoihin ja työn laadun mukaan alaryhmiin kuin henkisen työn panos.

1) Henkisen alan työlliset, vakinaisessa tai siihen verrattavassa työsuhteessa olevat ruumiillisen työn tekijät sekä maanviljelijöitä lukuun ottamatta itsenäiset yrittäjät, jotka eivät ole ansiotöissä omien töiden ohella, esim. kalastajat, metsästäjät, käsityölliset ja teollisuuden pienyrittäjät. Tällä ryhmällä ajankäyttötaulukon täyttäminen on yleensä varsin helppo tehtävä, koska sairaus- ja lomapäiviä lukuun ottamatta kaikki merkinnät tulevat samaan ajankäyttöryhmään tai ainakin muutamiiin harvoihin. Vakainainen maataloustyöntekijä edustanee tässä ryhmässä vaikeinta tapausta.

2) Maanviljelijät, jotka eivät käy ansiotyössä. Tämä varsin suuri ryhmä on jo huomattavasti vaikeammin haastateltavissa kuin edellinen, koska oman maatilatalouden työt lomakkeessa on melko yksityiskohtaisesti jaoteltu. Omassa metsässä suoritettujen erilaisten töiden, metsästyksen ja kalastukseen käytettyjen hajakäytävien yms. muistaminen tuottaa usein vaikeuksia.

3) Pelkkiä ansiotöitä tekevät tilapäistyöntekijät ovat edelliseen ryhmään verrattavia, joskus sitä helpompiakin haastateltavia. Lukuisat metsä-, uitto-, maa-, tie- yms. töitä tekevät työmiehet kuitenkin ovat usein maanviljelijöitä vaikeampia koehenkilöitä, varsinkin milloin verokirjaa ei ole käytettävissä. Heidän työvuotensa on hajanainen, työpaikkojen vaihdoksia voi olla 5—8, välillä vapaa-aikoja, matkoja ja työttömiäkin päiviä. Verokirjamerkinnöissä eivät pienet työn keskeytykset tavallisesti ilmene.

4) Kaikkein vaikeimmin tutkittavan ryhmän muodostavat kuitenkin osittain omaa tilaansa viljelevät, osittain tilapäisissä lyhytaikaisissa ansiotöissä käyvät miehet, joilla ajankäyttö voi hajautua 10:een, jopa useampaankin käyttöryhmään muutaman päivän merkinnöin. Kaikki edellä selostetuissa koehenkilöryhmissä esiintyvät vaikeudet kasautuvat tällaisissa haastatteluissa, joista tulee pitkiä sekä haastattelijan että haastateltavan kärsivällisyyttä koettelevia keskustelutilaisuuksia. Usein niissä turvaututaan toisten saman perheen jäsenten apuun. Onneksi nämä koehenkilöt ainakin tekijän kokemuksen mukaan suhtautuvat tämännäköiseen tiedusteluun rauhallisesti. Niinpä I esitutkimuksessa vain yhdessä tapauksessa koehenkilö (oikeammin tämän vaimo) kieltäytyi antamasta tietoja.

Ranuan koehenkilöt kuuluivat miltei kaikki ryhmään 4. Tässäkin ryhmässä muistamista suuresti helpottaa se, että pientilallisten elämässä etenkin metsä- ja uittotyömaalle lähtö ja sieltä paluu ovat vuosittain samoihin aikoihin toistuvia merkkitapauksia. Kun haastateltavan vaimo ja muut perheenjäsenet usein osallistuivat lomaketta täytettäessä käytyyn keskusteluun, sekin nähtävästi vähensi mittausvirheitä.

On mahdollista, että joltakin alueelta saadussa näytteessä on sitä enemmän mittausvirheitä, kuta suurempi osa koehenkilöistä kuuluu vaikeusryhmiin 3 ja 4 ja päinvastoin voidaan mittaustuloksia pitää sitä luotettavampina, kuta suurempi osa koehenkilöistä kuuluu ryhmiin 1 ja 2.

I esitutkimuksen x'_j ja \bar{p}_j -estimaatit laskettiin kausittain, yhteiskuntaryhmittäin ja kylittäin, I esitutkimuksen tuloksia verrattiin kausittain II esitutkimuksen tuloksiin, Norrbottenin tutkimuksen tuloksiin ja v:n 1941 Suomen maatalouslaskennan tuloksiin (HEIKINHEIMO 1950 b, Työtömyyskomitean ... 1950 sekä IHATSU 1953). Näissä vertailuissa ei ollut todettavissa niin suuria mittausharhoja, että ne olisivat tehneet tulosten osoittaman työpanoksen ja muun ajankäytön riippuvaisuuden eri tekijöistä epätodennäköisiksi. Päinvastoin näyttivät tulokset varsin hyvin vahvistavan muiden tietojen perusteella saatuja ennakkokäsityksiä.

Tulosten »järkevyys» muutamissa tutkituissa jakautumisissa ei kuitenkaan riitä osoittamaan, ettei mittausharhoja silti voi olla olemassa, vaikakaan niillä ehkä ei ole vaikutusta nimenomaan tutkittuihin ilmiöihin. Niinpä oman työn panoksen riippuvaisuus yhteiskuntaryhmästä (viljelmän suuruusluokasta) voi olla »oikean» suuntainen, vaikka sanottu työpanos onkin systemaattisesti liian pieni. Toisaalta johtuu tutkimuksen luonteesta, että toisissa ajankäyttöryhmissä täytyisi olla vastaavansuuruisen, mutta vastakkaissuuntainen mittausharha.

Ranualla suoritetun II esitutkimuksen (kirjanpitolukituksen) yksityiskohtaiset tulokset, joita aiemmin ei ole esitetty, suovat myös mahdollisuuden em. tavalla tutkia kirjanpitomenetelmän mahdollisia mittausharhoja. Samalla ne hyvin kuvaavat tutkittavan ilmiön luonnetta.

Taulukkoon 6/s. 48 sekä sen perusteella piirrettyyn kuvaan 2/s. 50 on merkitty v:n 1950 jokaisena arkipäivänä metsätalouden ja uiton ansiotöissä (1) sekä omissa maatilatalouden töissään (2) työskennelleiden koehenkilöiden lukumäärä sadanneksina kaikista II esitutkimuksen koehenkilöistä (133 miestä). Taulukkoon on merkitty myös sunnuntai-, pyhä-, vapun- ja itsenäisyyspäivinä näissä töissä olleiden miesten määrät, mutta selvyyden vuoksi ne on jätetty merkitsemättä kuvaan. Mainittujen luku-

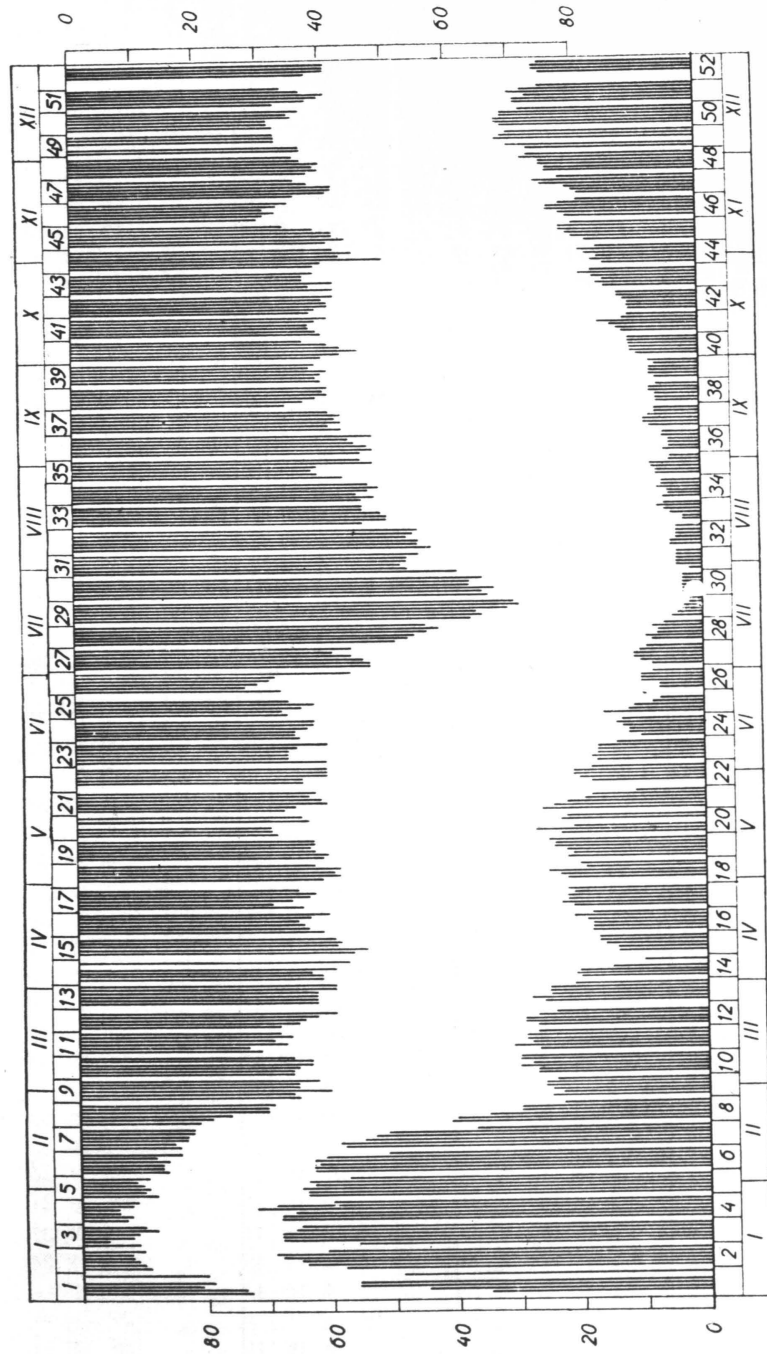
jen summa vähennettynä 100:sta (kuvassa pylväiden pystysuora väli) ilmaisee siten niiden koehenkilöiden suhteellisen määrän, jotka ovat ilmoittaneet työskennelleensä muissa töissä, olleensa opiskelemissa, asevelvollisuuttaan suorittamassa, sairaina, vapaapäivää viettämässä tai työttömyyden ja muiden pakottavien syiden takia työstä poissa.

Kuvaan 2/s. 50 merkityissä aikasarjoissa voidaan havaita odotettu kausivaihtelu ja sen lisäksi myös verrattain selvä viikon pituinen vaihtelu. Viikon alku- ja loppupäivinä oli metsä- ja uittotyömiesten määrä keskimäärin huomattavasti pienempi kuin viikon keskellä olevina päivinä. Omissa maatilatalouden töissä viikon sisäinen vaihtelu oli keskimäärin vastakkaissuuntainen. Tällainen ilmiö on odotettavissa, koska metsä- ja uittotyömaat sijaitsivat yleensä etäällä koehenkilöiden perheasunnoista. Niinpä puolet kaikista kirjatuista ansiotyöpäivistä oli sellaisia, jolloin koehenkilöt ilmoittivat yöpyneensä kotinsa ulkopuolella.

Mainitut aikasarjojen vaihteluilmiöt tuntuvat todenmukaisilta. Sarjoissa esiintyy kolmaskin vaihteluilmiö, jonka voi epäillä johtuvan mittausharhasta. Kuukauden viimeisen ja seuraavan kuukauden ensimmäisen päivän välillä osoittaa kuva joissakin tapauksissa verrattain suurta muutosta toisessa aikasarjassa tai molemmissakin. Erityisesti näin on laita kuukausien II—III, VI—VII, VII—VIII, VIII—IX ja X—XI vaihteessa. Kun huomattava osa kuukausien kirjanpitokalentereista todennäköisesti on täytetty muistinvaraisesti vasta kuukauden päätyttyä, on muistivirhe juuri ko. kuukauden alkupäivinä voinut tulla huomattavan suureksi. Osaksi virheet voivat olla satunnaisia, ja ne eivät aineiston pienuuden takia ole päässeet riittävästi tasoittumaan, mutta ne voivat olla myös systemaattisia.

Toisaalta saattaa olla mahdollista, että työmaita avataan tai lopetetaan juuri kuukauden vaihteessa. Elokuun alussa taulukosta havaittava metsätyöntekijöiden äkillinen lisäys ja saman kuun lopussa esiintyvä vähennys voisi siten olla todellinenkin. Myös maatilataloudessa lienee totuttu aloittamaan työt juuri kuukauden alussa tai päättämään ne kuukauden loppuun mennessä (esim. heinänteko heinäkuussa tai kotitarvepuun korjuu marraskuussa). Esitetyn tarkastelun tulos osoittaa, kuinka vaikeaa tutkimustulosten perusteella on tässäkin tapauksessa tehdä varmoja päätelmiä niiden harhaisuudesta.

Esitutkimuksissa kokonaisen vuoden ajankäytön haastattelussa mahdollisesti esiintyvistä mittausharhoista ei voitu tehdä ehdottoman varmoja päätelmiä. Kun työnantajilla ei ole olemassa luotettavia tietoja etenkin metsätyöntekijöittensä työajasta, ei Norrbottenissa käytettyä



Kuva 2. Metsätalouden ja uiton ansiotöissä (alempi pylväs) sekä omilla maatilatalouden töissä (ylempi pylväs) työskennelleiden miesten osuus sadanneksina koko miestyövoimasta v:n 1950 arkipäivinä. Ranua, II esitutkimus. Lähde: taulukko 6/s. 48.

Figure 2. Males employed in paid forest and floating work (lower column) and on their own farms (upper column) in per cent of total male labour force. Week-days in 1950. Ranua, Pilot Survey II. Source: Table 6/p. 48.

vertailumenettelyä voitu soveltaa. Ne mittausharhan tutkimukset, jotka voitiin suorittaa, vahvistivat luottamusta haastattelumenetelmän käytökelpoisuuteen. Tärkein haastattelututkimuksen mahdollisen harhaisuuden selvitys, II esitutkimuksen ja Ranualla v. 1951 suoritettujen haastattelututkimuksen tulosten vertailu, voitiin suorittaa vasta päätutkimuksen tulosten valmistuttua. Se selostetaan sen tähden luvussa 332.

Haastattelumenetelmään verrattuna osoittautui kirjanpitoimenetelmä ajankäytön mittauksessa monin verroin työläemmäksi ja kalliimmaksi. Vaikka II esitutkimuksen otos käsitti vain 133 koehenkilöä, sen loppuun saattaminen vaati useita tuhansia kirjeitä, kymmeniä aineiston kerääjiä, useita parin viikon kiertomatkoja tutkimusalueella sekä lisäksi ainakin yhden toimistotyövuoden. Samansuuruisen aineiston yksi haastattelija olisi kerännyt parissa viikossa. Kirjanpitoimenetelmän käyttäminen useita tuhansia koehenkilöitä käsittävässä otoksessa osoittautui siten mahdottomaksi käytännössä toteuttaa.

224. Haastattelulomake

Ensimmäisessä esitutkimuksessa käytetty haastattelulomake [liitteissä] oli vähäisiä eroja lukuun ottamatta samanlainen kuin Norrbottenissa käytetty (julkaisematon) lomake. Esitutkimuksissa havaittiin, että mitaustapaa oli mahdollisuus kehittää.

Norrbottenissa samoin kuin I esitutkimuksessa havaintovuosi jaettiin »luonnollisiin» työkausiin, joiden pituus oli $1\frac{1}{2}$ —3 kuukautta. Varsinkin III esitutkimus osoitti, että koehenkilön oli helpoin muistaa kalenterikuukauden ajankäyttö. Koko vuoden ajankäytön rakentaminen kalenterikuukausittain muistiin palautetuista luvuista oli huomattavasti helpompaa kuin kaikkien vuoden päivien jakautumisen suora arvioiminen.

Kun lomakkeessa kokeiltiin neljänneskuukauden (viikon) mittaus-tarkkuutta, havaittiin, että pienillä ajankäyttöryhmillä, kuten työttömyys- ja vapaa-ajalla, polttopuiden hakkuutyöpanoksella yms. oli taipumus pienetä ja tulla osittain sisällytetyiksi suurehkoihin työpanosryhmiin. Täten päivän mittaus-tarkkuus osoittautui parhaaksi, vaikka tulokset eivät sitä vaatineetkaan. Tarvittaessa on silti helppo soveltaa viikon tai 10 päivän tarkkuutta.

Norrbottenissa käytettiin mittayksikkönä päiviä, joihin myös sunnuntai- ja pyhäpäivät ja muut viralliset vapaapäivät sisältyivät. Täten saavutettiin se etu, että jonkin ajanjakson päivämäärästä voitiin suoraan laskea ajankäytön määrä. Siten esim. 1—21. 2. merkitsi 21 päivää. Muun-

taminen arkipäiviksi suoritettiin aineistoa käsiteltäessä vakiokertoimella (WIESLANDER 1948 s. 8). Esitutkimukset osoittivat, että haastattelu ei ainakaan mainittavasti ollut sen helpompi, käytettiinpä mittayksikkönä päivää tai arkipäivää. Pääasia oli, että lomakkeeseen oli merkitty kunkin kauden mittayksiköiden summa. Kirjanpilotutkimus osoitti (taulukko 6/s. 48), että metsä- ja uittotöitä tehdään usein myös virallisina vapaa-päivinä. Jotta tällaistenkin päivien työpanos tulisi huomioon otetuksi, tuli mainittu summa (a) arkipäivää mittayksikkönä käytettäessä tarvittaessa voida muuttaa suuremmaksi. Päivää mittayksikkönä käytettäessä mainittua työpanosta on hyvin vaikea ottaa huomioon.

Peitetyn työttömyyden, jota ruvettiin nimittämään kansanomaisesti *luppoajaksi*,¹ mittaaminen osoittautui vaikeaksi. Norrbottenissa (Anvisningar ... 1944, s. 12) luppoajalla (halvarbetslöshet) tarkoitettiin sitä aikaa, »jolloin koehenkilö ei ollut saavuttanut niin suurta työntuotosta, että sitä voitiin pitää täyden päiväpalkan arvoisena. Jos hän siten koko 46 päivän kautena oli tehnyt maatilallaan työtä, joka vastasi 30 koko päivän tehokasta ja hyödyllistä työtä, oli jäännös, so. tässä tapauksessa 16 päivää, merkittävä ko. kauteen luppoajaksi».²

Tällä tavalla mitattuna ei luppoaika enempää kuin se työpanoskaan, josta luppoaika on vähennetty, osoita enää, kuinka aika todellisuudessa on käytetty, vaan miten se olisi *haluttu* käyttää. I esitutkimuksen jälkeen siirryttiin uuteen luppoajan mittaamisen menetelmään. Työpanosluvut merkittiin sellaisinaan, samoin muu ajankäyttö, kuten avoin työttömyys. Täydentävänä tietona ajankäyttötietojen lisäksi merkittiin, kuinka monta oman työpanoksen päivää koehenkilö olisi voinut käyttää ansiotöihin, jos sellaisia aina olisi ollut saatavissa riittävän lähellä hänen asuinpaikkaansa. Kun oman työn tuotoksen raha-arvoa on koehenkilöiden varsin vaikea arvioida, tuli luppoajan määrittäminen näin ainakin jonkin verran helpommaksi. Kuitenkin osoittautui, että koehenkilöt eivät mielellään tunnustaneet luppoaikaa olevan olemassa. Seuraavantapainen lausunto oli tyypillinen: »Eihän näiltä omilta töiltään joutaisi ollenkaan ansiotöihin, mutta kun on saatava rahaa perheen elatukseen». Koehenkilöiden oli vaikea kuvitella, että kevättalvella voisi olla riittävästi ansiotöitä. Luppoaika on, mitattakoon se miten tahansa, epämääräinen suure ja suuresti

¹ Kansankielen sanakirjan kokoelmissa on useita merkintöjä luppoaika-termistä. Esim. Kärsämäki 1933: »Kiireellisten töitten väliaika, jolloin talonväellä ei ollut sanottavasti töitä. Ajan kuluksi kapisteltiin milloin mitäkin, mm. kalastusta harjoitettiin kevät- ja kesäkiireitten välisenä luppoaikana.» (Sanakirjasäätiöltä saatu ennakkotieto).

² Suomenkos tekijän.

Taulukko 7. Oman työn panoksen eri alaryhmiin sisältyvän luppoajan määrä sadanneksina luppoajan koko määrästä sekä ryhmän koko työpanoksesta. Ranua. II esitutkimus.

Table 7. Underemployment in different subgroups of unpaid labour input, in per cent of total underemployment and of total labour input in each subgroup. Ranua. Pilot Survey II.

Oman työn panoksen alaryhmä Subgroup of unpaid labour input	Eri alaryhmiin sisältyvä luppoaika sadanneksina Underemployment in the different subgroups in per cent of	
	luppoajan koko määrästä total amount of under-employment	ryhmän koko työpanoksesta total labour input in the group
1. Talonrakennus ja uudisraivaus — <i>House construction and land clearing</i>	8	1
2. Muu maatalous — <i>Other agricultural work</i>	44	5
3. Kotitarvepuun hakkuu ja ajo — <i>Felling and hauling of wood for domestic use</i>	39	10
4. Markkinapuun hakkuu ja ajo — <i>Felling and hauling of wood for sale</i>	5	2
5. Metsästys, kalastus ym. — <i>Hunting, fishing etc.</i>	4	6
6. Oman työn panos yhteensä — <i>Total input of unpaid labour</i>	100	5

riippuvainen haastattelijan kyselytavasta. Etenkin Suomen kaltaisessa maassa on kuitenkin tärkeää saada se arvioiduksi vaikkapa likimäärin.

II esitutkimuksessa voitiin myös todeta, millä omilla töillä luppoaika oli peitetty. Tulos näkyy taulukosta 7. Valtaosa (83 %) luppoajasta oli peitetty maatalouden töillä tai polttopuiden korjuulla (viimeksi mainittuun ryhmään II esitutkimuksessa sisällytettiin myös polttopuiden pilkkominen). Koehenkilöiden arvion mukaan polttopuiden korjuuseen käytettyä työtä olisi ollut mahdollisuus suhteellisesti eniten tehostaa. He pitivät 10 % tästä työpanoksesta luppoaikana.

Norrbottenissa käytetty ajankäytön kyselyjärjestys: ansiotyöt, omat työt ja muu ajankäyttö osoittautui muuten hyväksi, mutta viimeksi mainittu ryhmä näytti olevan paras sijoittaa kyselyssä ensimmäiseksi. Luppoaikaa koskevat kysymykset tuli tehdä viimeiseksi.

3. Päättökäsimus. Vuoden 1950 maaseudun työvoiman tutkimus

31. Päättökäsimuksen tarkoitus ja rajoitus

Valtakunnan työvoimaa koskevat tiedot tulevat tarpeellisiksi silloin, kun ryhdytään toimenpiteisiin talouselämän ja sosiaalisten olojen kehittämiseksi. Niiden suunnitelmat vaativat joukon perustietoja, jotka ovat saatavissa vain erityisillä tutkimuksilla. Toisaalta näiden toimenpiteiden kohde ja laatu määräävät, millaisia asioita on tutkittava ja miten.

Varsinkin toisen maailmansodan jälkeen on Suomessa tehty pitkälle meneviä suunnitelmia taloudellisen toiminnan monipuolistamiseksi ja edistämiseksi sekä yhteiskunnallisten olojen parantamiseksi. Mainittakoon teollistamissuunnitelmat (Maaseudun elinkeinokomitean mietintö 1951, Teollistamiskomitean mietintö 1951, KEKKONEN 1952), metsätalouden kehittämistä tarkoittavat suunnitelmat (Puun maakuljetusolojen kehittäminen 1952, OSARA 1953), maatalouden uudelleenjärjestelyyn tähtäävät aloitteet (ELLILÄ 1953, JÄNTTI 1953) sekä valtakunnalliset yleissuunnitelmat (Valtakunnansuunnittelukomitean mietintö 1954).

Jos suurisuuntaisiin investointeihin tai muihin suuria työvoimamääriä vaativiin toimenpiteisiin ryhdytään eikä ole välttämätöntä vapaata työvoimaa, voi tästä olla tuhoisat seuraukset koko maan talouselämälle, vaikka ulkomaisin lainoin tai muuten olisikin saatu tarpeelliset varat. Hyvänä esimerkkinä on toisen maailmansodan aikana ja sen jälkeen useissa maissa ylityöllisyyden (ja muiden tekijöiden) aiheuttama inflaatiopaine, joka pakotti rajoittamaan säännöstelyllä talouselämän aktiviteettia.

Toisaalta on myönnettävä, että vaikka työvoimavarojen määrä tunnetaan, sen perusteella tehdyt laskelmat niin sanottujen työvoimareservien määristä ja kehityksestä eivät läheskään aina pidä paikkaansa. Siihen vaikuttavat monet ihmisen käyttäytymistä säätelevät tekijät, kuten KEYNESIN (1951, s. 126) korostama kulutusalttiuden jäykkäliikkeisyys. Tästä huolimatta laskelmilla voi olla tärkeä tehtävä talous- ja sosiaalipolitiikan ohjaajina.

Kansainvälinen työtoimisto (Employment ... 1948, ss. 6—9) pitää

työvoimaa ja työllisyyttä koskevia tietoja tarpeellisina miltei jokaiselle paikalliselle ja yleiselle talous- ja sosiaalipoliittiselle toimenpiteelle, kuten verotus-, maatalous- ja koulupolitiikalle sekä sosiaalivakuutukselle. Kansakunnan työvoimaa ja sen käyttöä koskevat tiedot, kuten muidenkin tuotannon tekijöiden ja niiden käytön tuntemus ovat tarpeellisia myös muista syistä. Usein talouselämän puutteellisuudet tulevat esille vasta tutkimuksissa, ja yhtä usein pelkkä tieto olosuhteista riittää ohjaamaan kehitystä ilman erityisiä valtion toimenpiteitä.

Esitutkimusten ollessa vielä keskeneräiset Tilastollisen päätoimiston ylijohtajan professori MARTTI KOVERON johdolla asetettu toimikunta päätti, että vuoden 1950 väestönlaskennan yhteydessä oli kerättävä aineisto myös työvoimatutkimusta varten. Näin päätetty erikoistutkimus sai nimen *vuoden 1950 maaseudun työvoiman tutkimus*.

Maaseudun työvoiman tutkimuksen päämäärä oli, kuten luvussa 14 selostettiin, kahdenlainen. Toisaalta pyrittiin selvittämään työvoiman kokonaismäärä, ajankäyttö ja työpanoksen elinkeinorakenne sekä kaikkien edellä mainittujen kausivaihtelu v. 1950. Toisaalta haluttiin eräiden ammattiryhmien, ennen kaikkea metsätaloudessa ja uitossa, mutta samalla myös maataloudessa toimivan väestönosan suuruutta, rakennetta ja sosiaalisia oloja koskevia tietoja. Niiden tarkoitus oli täydentää väestönlaskennassa saatavia, kausityöllisten osalta pakosta epätäydellisiksi jääviä tuloksia ammattiryhmityksestä ja eri ammattiryhmien yhteiskunnallisista oloista. Työvoimatutkimusten tulisi kansainvälisten ohjeiden mukaan peittää koko valtakunnan työvoima. Olisi ollut luonnollista, että päättökäsimusta varten olisi poimittu näyte koko maan väestöstä tiettyyn alimmaisikärajaan saakka. Vuoden 1950 yleisen väestön- ja maatalouslaskennan kustannukset olivat kuitenkin valtiolle niin suuri menoerä, että ensimmäinen ja laadultaan vielä kokeilunluonteinen työvoimatutkimus voitiin suorittaa vain välttämättömimmät tarpeet tyydyttävässä laajuudessa. Tästä syystä tutkimus rajoitettiin maantieteellisesti vain maalaiskunnat peittäväksi. Tutkimusta suunniteltaessa voitiin teollisuuden ja muiden kaupunkielinkeinojen työvoima ja sen ajankäyttö jättää erittelemättä ja kohdistaa päähuomio maaseudun kausielinkeinoihin, joiden työvoimakysymys sekä kansantaloudellisesti että sosiaalipoliittisesti kaippaa eniten valaisua.

Naistyövoima jätettiin tutkimatta. Maataloutta lukuun ottamatta naistyövoimalla tutkimuksen pääkohteena olevissa maaseutuelinkeinoissa on varsin vaatimaton merkitys. Maatalouden työvoimakysymys sen sijaan olisi edellyttänyt naisten työpanoksen tuntemista. Kotitalouden ja maa-

talouden työpanoksen erottaminen naistyövoiman osalta on kuitenkin vaikea, usein mahdoton tehtävä haastattelututkimuksessa. Kun naistyövoiman mukaan ottamisella kustannukset olisivat nousseet miltei kaksinkertaisiksi ja kun maatalouslaskennan työvoima-aineisto peitti valtaosan mautilojen naistyövoimasta, katsottiin parhaaksi rajoittaa tutkimus pelkästään miehiin.

Miespuolisesta väestöstä erotettiin tutkimuksen ulkopuolelle ammattisotilaat ja laitoshoidokit (Employment ... 1948, s. 119). Lisäksi alle 15-vuotiaat (vuoden 1935 jälkeen syntyneet) sekä yli 64-vuotiaat (ennen vuotta 1886 syntyneet) päätettiin jättää tutkimatta. Työkykyisessä iässä olevana väestönä pidetään yleisesti ikäluokkia 15—64 vuotta (HYPPÖLÄ, TUNKELO ja TÖRNQVIST 1949, s. 34). Kuitenkin on kyseenalaista, asetettiiniko sanotut ikäraajat maaseudun työvoiman tutkimuksessa oikein. USA:n työvoiman tutkimuksissa määritellään vain alin ikäraja (14-vuotiaat luetaan työvoimaan), ja JAFFE ja STEWART (1951, s. 43) katsovat, että tätä ikärajaa olisi huomattavasti alennettava vähemmän teollistuneiden maiden työvoiman tutkimuksissa. Alimman ikärajan määrää nuorten vuosiluokkien työhönastumisikä. Otantatutkimuksissa ei näytettä kannata ottaa sellaisesta ikäluokasta, josta vain mitättömän pieni osa kuuluu työvoimaan.

Väestönlaskennassa luettiin ammatissa toimivaan väestöön 13 vuotta vanhemmat ammatin harjoittajat. Taulukossa 8 esitetään 14-vuotiaiden ja yli 64-vuotiaiden suhteellinen osuus lääneittäin maalaiskuntien miespuolisesta ammatissa toimivasta väestöstä.

Väestönlaskennan mukaan oli yli 64-vuotiaita ammatissa toimivia miehiä 6.9 % koko maaseudun ammatissa toimivasta miespuolisesta väestöstä. Maaseudun työvoiman tutkimuksen näytteelle asetettu ylin ikäraja oli ilmeisesti virheellinen ja ehkä kokonaan tarpeeton. Sen sijaan näyttää taulukon 8 perusteella, että 14-vuotiaiden mukaan ottaminen olisi vain vähäisessä määrin voinut lisätä tulosten käyttökelpoisuutta. Kuitenkin on todennäköistä, että vaikka väestönlaskennassa 14-vuotiaat pojat vain harvoissa tapauksissa katsottiin ammatin harjoittajiksi, he paljon useammin ovat osan vuodesta tehneet etenkin maa- ja metsätalouden töitä. Maaseudun työvoiman tutkimuksen tulosten mukaan maaseudun 15-vuotiaista pojista 33 % teki v. 1950 ainakin jonkin päivän ansiotyötä metsätaloudessa ja uifossa. Ikäluokassa 60—64 vuotta mainittu suhdeluku oli 31 %. Tästä päätellen näinkin raskaaseen työhön osallistuu huomattava määrä alle 15-vuotiaita poikia samoin kuin yli 64-vuotiaita miehiä. Maatalouslaskennan tuloksista (Yleinen maatalouslaskenta 1954) tekijä on

Taulukko 8. Maaseudun 14-vuotiaiden sekä 65-vuotiaiden tai sitä vanhempien ammatissa toimivien poikien ja miesten osuus maaseudun miespuolisesta ammatissa toimivasta väestöstä v. 1950. Väestönlaskennan mukaan.

Table 8. Gainful male workers of 14, and of 65 years and over in per cent of all gainful male workers. All rural districts, 1950. According to the Population Census of 1950.

Lääni County	Ammatissa toimiva miespuolinen väestö yhteensä Total number of gainful male workers	siitä sadanneksina — of which in per cent	
		14-vuotiaat males of 14	65-vuotiaat tai vanhemmat males of 65 or over
Uudenmaan — Uusimaa	67 488	0.3	12.3
Turun ja Porin — Turku and Pori	132 790	0.3	7.1
Ahvenanmaa — Åland	5 791	0.2	7.0
Hämeen — Häme	95 617	0.3	8.0
Kymen — Kymi	59 984	0.3	7.0
Mikkelin — Mikkelä	60 702	0.4	7.1
Kuopion — Kuopio	119 108	0.4	5.7
Vaasan — Vaasa	144 426	0.5	5.7
Oulun — Oulu	90 518	0.6	5.7
Lapin — Lapland	37 232	0.3	5.7
Koko maaseutu — All rural districts ..	813 656	0.4	6.9

laskenut, että alle 15-vuotiaiden poikien ja tyttöjen työpanos maatilatalouden koko työpanoksesta oli 4 %, jos mittayksikkönä käytetään muuntaumatonta työpäivää.

Työvoimatutkimuksia vastaisuudessa suoritettaessa on syytä peittää niillä nuoret ikäluokat ainakin 14-vuotiaista alkaen ja kaikki sitä vanhemmat ikäluokat.

32. Päättökäytännön menetelmä

321. Perusaineiston keräystapa ja suuruus

3211. Poimintamenetelmä

Maaseudun työvoiman tutkimuksen aineisto kerättiin henkilökohtaisella ajankäytön kyselymenetelmällä näytteestä, joka systemaattista valintaa käyttäen poimittiin perusjoukosta. Perusjoukon muodostivat luvussa 3213 mainittuja poikkeuksia lukuun ottamatta kaikki maan maalaiskun-

nissa 31. 12. 1949 asuneet vv. 1886—1935 syntyneet miespuoliset henkilöt. Näistä pyrittiin poimimaan näytteeseen kaikki jonkin, minkä tahansa, kuukauden 1. tai 15. päivänä syntyneet miehet ja pojat, joita seuraavassa nimetään T-mieheksi.

Tutkimuksen suunnitelman mukaan näytteenottomenetelmälle asetettiin seuraavat neljä vaatimusta. Näytteen tuli olla otos tai ainakin sitä vastaava näyte tutkittavasta perusjoukosta. Sen oli oltava niin suuri, että tutkimuksen päätulokset voitaisiin laskea riittävän varmoina ainakin metsänhoitolautakuntien toiminta-alueita vastaaville maantieteellisille alaryhmille. Koehenkilöt tuli voida valita väestönlaskennan laskentatyön yhteydessä myös sellaisissa tapauksissa, jolloin heitä ei oltu merkitty käytettävissä olleeseen luetteloon (henkikirjan otteeseen). Lopuksi tuli näytteestä lasketut arviot alustavasti voida suurentaa koko perusjoukkoa vastaaviksi perusjoukon suuruutta tuntematta.

Näyte täyttää otoksen vaatimukset, mikäli jokaiselle näyteyksikölle on ennen valintaa määrätty toisista näyteyksiköistä riippumaton todennäköisyys tulla näytteeseen (esim. The preparation ... 1950, s. 9). Systemaattinen näyte ei yleensä täytä tätä vaatimusta, koska yksiköiden todennäköisyys tulla valituksi riippuu toinen toisistaan. Sen tähden systemaattisen näytteen käyttö tutkimuksissa otoksen sijasta jää kokonaan tutkimuksen suorittajan vastuulle (YATES 1949, s. 30).

Tästä huolimatta systemaattista valintaa on usein käytetty sen monien etujen vuoksi. Päätutkimuksessa ei ollut saatavissa kehikkoa, josta myös kuntiin v. 1950 muuttaneiden näyteyksiköiden valinta olisi voitu suorittaa etukäteen arpomalla, eikä tätä arpomistakaan olisi käytännöllisistä syistä voitu jättää kenttähenkilökunnan tehtäväksi. Näin ollen jäi ainoaksi mahdollisuudeksi valita näyteyksiköille tunnus, joka oli vaivattomasti todettavissa milloin tahansa laskija tapasi perusjoukkoon kuuluvan henkilön.

Ruotsin väestönlaskennassa v. 1945 niin sanottuun kahdestoistaosana näytteeseen poimittiin perusjoukon heinäkuussa syntyneet henkilöt (Folkräkningen ... 1948, ss. 2*—3*). Syntymääjan ottaminen valintaperusteeksi maaseudun työvoiman tutkimuksessa tuntui varsin edulliselta, koska saattoi olettaa, että tämä tieto oli käytännöllisesti katsoen jokaiselta kyseeseen tulevalta henkilöltä vaivattomasti saatavissa. Määrätyn kuukauden valinnasta luovuttiin seuraavasta syystä. Syntyvytydessä on havaittavissa Suomessa selvä kausivaihtelu (Suomen tilastollinen vuosikirja). Vaikka tämän kausivaihtelun määrää alueittain ja yhteiskuntaryhmittäin ei meillä ole tietävästi tutkittu, on mahdollista, että se on erilainen, ehkäpä erisuuntainenkin maan eri alueilla ja eri ammattiryhmiin

kuuluissa perheissä. Niinpä on oletettavissa, että niissä Itä- ja Pohjois-Suomen perheissä, joissa perheenisät varsinkin entisaikoina säännöllisesti viettivät talvikuukaudet kaukaisilla metsätyömailla, syntyvyys joulutammikuun tienoilla oli keskimäärin suurempi kuin muina kuukausina. Näin ollen on mahdollista, että metsätyömiesten lapsilla on pienempi todennäköisyys tulla mainitulla tavalla poimittuun näytteeseen kuin muilla. Siten voidaan epäillä, saadaanko yhtä kuukautta tunnuslukuna käyttäen harhaton näyte.

Mainittujen vaikeuksien välttämiseksi päätettiin syntymäpäivälukuja ottaa kaikista kuukausista. Esitutkimusten perusteella arvioitiin, että kuukauden kaksi päivää antaa riittävän näytteen. Koska päivämäärien 1—28 katsottiin olevan samanarvoiset, valittiin tunnusluvuiksi helposti muistettavat »pyöreät» luvut 1 ja 15, jolloin ko. päivämäärien etäisyys toisistaan tuli samalla olemaan suunnilleen tasavälinen.

Täten vain niillä perusjoukon henkilöillä, joiden syntymäpäivä sattui olemaan kuukauden 1. tai 15. päivä, oli mahdollisuus tulla näytteeseen. Kuitenkin voidaan esittää se vaikeasti kumottava väite, että naisen munasolulla on yhtä suuri todennäköisyys hedelmöityä sellaisella hetkellä, että se johtaa kuukauden 1. tai 15. päivänä tapahtuvaan synnytykseen, kuin sellaisella hetkellä, että se johtaa synnytykseen minä tahansa muuna kuukauden 28:sta ensimmäisestä päivästä. Ikivanha ihmisiin juurtunut uskomus kosmisten sekä ihmisen fyysillisten ja psyykkisten ilmiöiden yhteenkuuluvaisuudesta on tosin sitkeästi ylläpitänyt tutkijoissakin käsitystä, että taivaankappaleiden kiertoliikkeiden ja naisen menstruaatio-syklin välillä olisi korrelaatio ja että tämän seurauksena synnytysten lukuisuus vaihtelisi n. 30 päivän jaksoin. HOSEMANN (1950) on voinut osoittaa, että tällaiset tutkimustulokset miltei poikkeuksetta ovat virheellisiä.

Siten ei ole olemassa syytä, minkä takia maalaisväestöstä otettu systemaattinen näyte, jossa näyteyksikön valinnassa käytetään satunnaislukuna henkilön todellista syntymäpäivää, ei olisi käytännössä otosta vastaava.

Samalla on osoitettu, että sellaisen (täydellisen) näytteen suurennustekijä teoriassa saadaan jakamalla perusjoukon kaikkien potentiaalisten syntymäpäivien lukumäärä luvun 12 niin monennella kerrannaisella kuin tunnuslukuja kuukaudesta on otettu. Kyseessä olevassa tapauksessa teoreettinen suurennustekijä on $365.25/24$ (vuoden 1900 puuttuva karkauspäivä on tässä jätetty huomioon ottamatta).

3212. Mittausmenetelmä

Kustakin luvussa 3213 määritellystä T-miehestä tuli väestönlaskijan asianomaista koehenkilöä haastatteleamalla täyttää ns. T-lomake, jonka eräs täytetty kappale on jäljennetty liitteissä.

T-lomakkeen 1. sivulle kirjoitettiin tietojen tarkistuksessa välttämättömien koehenkilön identifioimismerkintöjen lisäksi seuraavat tiedot (numerot viittaavat lomakkeen ao. kohtiin).

2) Asuinkunta 31. 12. 1950, jonka perusteella koehenkilön sijoitus määrättiin aineiston maantieteellisessä jaotuksessa.

5) Syntymäaika, päivä, kuukausi ja vuosi, joiden perusteella tarkistettiin, oliko haastateltu henkilö T-mies, sekä määrättiin hänen sijoituksensa ikäluokituksessa.

7) Sen tilan tai tontin pelto-, puutarha- ja metsäpinta-ala, jota koehenkilö v:n 1950 aikana viimeksi omaan lukuunsa viljeli tai, jollei hän viljellyt viljelmää v:n 1950 aikana, vastaavat tiedot siitä tilasta tai tontista, jolla hän asui 31. 12. 1950.

8) Koehenkilön hallintasuhte mainittuun tilaan tai tonttiin. Kahden viimeksi mainitun tiedon tarkoituksena oli aineiston jakaminen yhteiskuntaryhmiin. Samoin tarkoituksena oli tutkia koehenkilöiden ajankäytön rakenteen suhdetta viljellyn pellon ja puutarhan sekä hallinnassa olevan metsämaan pinta-aloihin. Lisäksi tuli näiden tietojen avulla tutkia myös viljelmillä sekä maatilametsälöillä hehtaaria kohti käytettyä työpanosta ja sen suhdetta viljelmien ja metsälöiden pinta-aloihin.¹

9) Koehenkilön työsuhteen pysyvyys pyrittiin selvittämään kahdella tähän kohtaan merkityllä kysymyksellä. Tarkoituksena oli tämän vastauksen perusteella jakaa toisen paikallisessa palveluksessa olleet koehenkilöt vakinaisen työsuhteen omaaviin ja tilapäisiin työntekijöihin. Kysymysten asettelu ei ollut täysin mainittua tarkoitusta vastaava, joten tämän kohdan vastausten merkitys oli etupäässä ajankäyttötaulukkoa kontrolloiva.

10) Työkykyisyyden mittapuuksi asetettiin monien vaihtoehtojen jälkeinen kysymys: »Oliko sairauden tai invaliditeetin takia vuonna 1950 tekemänne työn määrä pienempi kuin puolet siitä, minkä olisitte samaa työtä tehnyt terveenä?» Myönteinen vastaus tähän kysymykseen johti T-miehen

¹ 8. kohdan kysymys osoittautui käytännössä liian vaikeatajuiseksi. Sen oikeampi muoto olisi ollut: »Oletteko tämän viljelmän tai tontin haltijaperheen kotona asuva jäsen?»

siirtämiseen pääaineistosta erikseen käsiteltävään työkyvyttömiin ryhmään. Minkäänlainen lääketieteellinen työkykyisyyden mittaus ei väestönlaskennassa ollut mahdollista. Kuitenkin oli välttämätöntä erottaa aineistossa sellaiset koehenkilöt, joiden työttömyyden, sairauden, vapaa-ajan ym. muuhun kuin työhön käytettyjen päivien määrä oli työkyvyttömyyden takia luonnottoman suuri ja työnsuorituskyky myös työpäivinä hyvin pieni. Asetettu kysymys lienee tässä täyttänyt tarkoituksensa, kun mainittujen koehenkilöiden lisäksi myös ne, joilla sairauspäiviä oli enemmän kuin 150, siirrettiin työkyvyttömiin joukkoon. Tosin myös osa niistä esim. tapaturmaisesti v. 1950 loukkaantuneista T-miehistä, joiden työnsuorituskyky heidän terveinä ollessaan v:n 1950 aikana oli ollut vähentymätön, tuli pääaineistosta karsituksi ja heidän työpanoksensa päätuloksissa jäi huomioon ottamatta. On kuitenkin todennäköistä, että heidän lukumääränsä koko aineistoon verrattuna oli häviävän pieni.

T-lomakkeen pääosan muodosti sen 2. ja 3. sivun ajankäyttötaulukko, johon laskijan oli merkittävä päivien määrää osoittavin numeroin T-miehen arkiajan käyttö vuonna 1950 kuukausittain jaotellen. Taulukon muoto perustui niihin kokemuksiin, joita esitutkimuksissa vv. 1949—50 oli saatu. Koska ajankäytön lajien valinnassa pyrittiin täyttämään eri tahoilta esitettyjä toivomuksia, tuli niitä varsin runsaasti, yhteensä 28. Näiden lisäksi lomakkeeseen sijoitettiin tila oman työn panokseen sisältyvien loppoaikapäivien sekä ansiotyöpanokseen sisältyvien työttömyys-työpäivien merkintää varten.

Ajankäytön mittauksen tuli laskijoille annettujen ohjeiden mukaan tapahtua siten, että laskija henkilökohtaisesti haastatteli koehenkilöä. Esitutkimuksissa kehitettyä kyselyjärjestystä noudattaen hän merkitsi, kuukauden kerrallaan, lomakkeeseen painettujen arkipäivien summan jakautumisen eri ajankäyttöryhmiin. Muistin helpottamiseksi tuli mikäli mahdollista käyttää verokirjaa apuna. T-miesten valmentamiseksi ennakolta oli tarkoitus ennen väestönlaskennan toimeenpanoa jakaa heille erityinen vuoden 1950 kalenteri, johon he olisivat voineet merkitä arkipäivien jakautumisen ansiotöihin, omiin töihin sekä muuhun ajankäyttöön. Kirjapainolakon tähden tätä toimenpidettä ei voitu suorittaa. Tutkimuksen perusteena oleva mittaus jouduttiin siten suorittamaan haastatteluhetkellä koehenkilön ja hänen perheensä muistiin palautuvien seikkojen ja verokirjan avulla, milloin se oli saatavissa.

3213. Näytteen peittävyys

Tutkimuksen näytettä poimittaessa selvin menettelytapa olisi ollut täyttää T-lomake rajoituksitta jokaisesta T-miehestä. Väestönlaskennan teknillinen järjestely asetti tälle esteitä, ja niin ollen laskentatoimikunnille annetuissa ohjeissa määrättiin T-lomake täytettäväksi seuraavista henkilöistä:

1) Kaikista niistä T-miehistä, joista täytettiin väestönlaskennan ruokakuntalomake¹ C, myös tilapäisesti ruokakunnassa läsnäolevista ja tilapäisesti poissaolevista henkilöistä.

2) Kaikista niistä T-miehistä, jotka laskentahetkellä asuivat metsä-, tie- tai muilla työmailla ja joista sen tähden täytettiin D-lomake,¹ useimmiten työnjohtajan toimesta.

Näiden ohjeiden mukaan T-lomaketta ei täytetty sellaisista T-miehistä, jotka laskentahetkellä olivat kunnallis-, vanhain- tai lastenkodissa, rangaistus- tai työlaitoksessa, koulukodissa, sairaalassa tai parantolassa tai muussa sellaisessa laitoksessa.

T-lomaketta ei myöskään tarvinnut täyttää puolustuslaitoksen toimesta sen varusmiehistä eikä majoitusliikkeiden ja alusten henkilökunnan toimesta hotelleissa, matkustajakodeissa, yö- ja matkailumajoissa ja Suomen satamissa tai aluevesillä aluksissa laskentahetkellä oleskelevista miehistä.

Kaikista edellä mainituista T-miehistä tuli täyttää T-lomake niissä tapauksissa, jolloin heidän voitiin katsoa kuuluvan sellaiseen ruokakuntaan, josta C-lomake oli täytettävä ja josta he vain tilapäisesti olivat poissa. Tämä määräys merkitsi, että joitakin harvalukuisia poikkeuksia lukuun ottamatta T-lomake piti täyttää tilapäisesti sairaalassa, vankilassa tai muussa laitoksessa olleista, varusmiehistä, hotelleissa ym. tilapäismajoituksessa laskentahetkellä asuneista ja myös kaikista niistä merimiehistä, jotka kuuluivat johonkin ruokakuntaan, mutta olivat laskentahetkellä merillä.

¹ Väestönlaskennassa täytettiin C-lomake jokaisesta ruokakunnasta ja jokaisesta asuinhuoneistosta: D-lomake täytettiin laskentahetkellä hotelleissa, matkustajakodeissa, metsätyömailla jne. olevista henkilöistä, varusmiehistä, sairaaloiden, kunnalliskotien yms. laitosten hoidokeista sekä vankiloissa, työlaitoksissa yms. säilytettävistä henkilöistä (Laskijan opas 1950, s. 10).

Lomakkeita käsiteltäessä aineistosta poistettiin

- 1) kaksoiskappaleet,
- 2) kaupunki- ja kauppalakunnissa (Karhulan kauppala lukuun ottamatta) laskentahetkellä asuneiden,
- 3) laitoksissa koko vuoden 1950 olleiden ja
- 4) puolustuslaitoksen ja rajavartioston vakinaisessa palveluksessa laskentahetkellä olleiden T-miesten lomakkeet.

Kun laskentateknillisistä syistä Karhulan kauppala kerättyä aineistoa ei voitu erottaa Kymen kunnan aineistosta, tuli koko aineisto maantieteellisesti peittämään kaikki maassa 1. 1. 1951 olleet 484 maalaiskuntaa ja sen lisäksi Karhulan kauppalan Kymen kuntaan sisältyvänä. Muut vv:n 1950/51 vaihteessa perustetut kauppalat jätettiin tutkimatta.

3214. Näytteen suuruus

Väestönlaskentatoimikunnat palauttivat täytettyjä T-lomakkeita yhteensä 47 895 kpl. Edellä mainituista syistä jouduttiin lopullisesti hylkäämään 1 042 lomaketta, joten näytteen suuruudeksi jäi 46 853 koehenkilöä. Lääneittäin oli koehenkilöitä seuraavasti:

Uudenmaan lääni	3 557	koehenkilöä
Turun ja Porin lääni	7 546	»
Ahvenanmaan maakunta	319	»
Hämeen lääni	5 302	»
Kymen lääni	3 645	»
Mikkelin lääni	3 562	»
Kuopion lääni	7 158	»
Vaasan lääni	8 294	»
Oulun lääni	5 354	»
Lapin lääni	2 116	»

Koko maaseutu 46 853 koehenkilöä

Hylätyt lomakkeet jakautuivat hylkäämisperusteen mukaan seuraavasti:

Syntymääjan perusteella näytteeseen kuulumattomia	332	lomaketta
Asuinkunta kaupunki tai kaupala	129	»
T-mies puolustuslaitoksen tai rajavartioston ammattisotilas	174	»
T-mies sairaalassa, muussa laitoshoidossa tai lomake epäselvä	306	»
Kaksoiskappaleita	82	»
Kuntiin jäi tarkistuksessa	13	»
Muu hylkäämisperuste	6	»

Yhteensä hylättiin 1 042 lomaketta

Kuten luvussa 3311 esitetään, eivät mainitut koehenkilöt käsitä koko ohjeiden mukaan poimittavaksi tarkoitettua näytettä, koska eri syistä ei täytetty T-lomaketta osasta haastateltaviksi tarkoitettuja koehenkilöitä.

322. Aineiston käsittely

3221. Lomakkeiden tarkistus ja kodifiointi

T-lomakkeiden käsittely voitiin aloittaa helmikuussa 1951, jolloin las-kentatoimikunnat rupesivat lähettämään niitä Helsingin yliopiston metsä-taloustieteelliselle laitokselle. Sen jälkeen kun lomakkeista oli poistettu näytteeseen kuulumattomat, suoritettiin niissä kolmivaiheinen tarkistus: ensimmäinen ja toinen esitarkistus sekä jälkitarkistus.

Ensimmäisen esitarkistuksen tehtävänä oli verrata lomakkeiden eri tie-toja toisiinsa, suorittaa niissä eräitä korjauksia ja osa kodifioinnista sekä erottaa puutteellisia tai ristiriitaisia tietoja sisältävät lomakkeet jälki-tarkistukseen. Käsittelyn yhdenmukaistamiseksi sitä suorittava henkilö-kunta pyrittiin supistamaan mahdollisimman pieneksi. Ensimmäisen esi-tarkistuksen suoritti vähäisiä poikkeuksia lukuun ottamatta seitsemän tarkastajaa, joista kolme oli metsänhoitajaa ja loput ylioppilaita. Esi-tarkastajille annetut tärkeimmät ohjeet ovat liitteissä.

Ensimmäisessä esitarkistuksessa hyväksytyt ja hylätyt (mutta eivät jälkitarkistukseen siirretyt) lomakkeet tarkisti tekijän ohjeiden mukaan toisessa esitarkistuksessa uudelleen metsänhoitaja AIMO BLOMBERG, joka siten tuli läpikäyneeksi valtaosan aineistosta. Hänen tehtävänä oli sakkomääräyksiin pitää ensimmäisen vaiheen esitarkastajien virhemerkintät vähäisinä, korjata tehdyt virheet ja huolehtia muutosmerkintöjen yhdenmukaisuudesta tekijältä työn kuluessa saamiensa ohjeiden mukai-
sesti. Ajankäyttöryhmät tulivat tässä käsittelyssä saamaan lomakkeeseen merkittyjä määritelmiä täsmällisemmän sisällön, joka selostetaan tutki-muksen päätulosten yhteydessä (HEIKINHEIMO 1954 b).

Jälkitarkistukseen ensimmäisestä esitarkistuksesta siirretyt puutteelli-set ja epäselvät lomakkeet käsitteli samoin tekijältä saamiensa ohjeiden mukaan valtiotieteiden kandidaatti TOINI RISTIMÄKI. Lukuun ottamatta verrattain vähäistä osaa, joka uuden harkinnan jälkeen voitiin hyväksyä, nämä lomakkeet (n. 15 % aineistosta) palautettiin kuntien väestönlas-kentatoimikunnille täydennyksiä ja korjauksia varten. Niistä ne palasivat uudelleen jälkitarkistukseen ennen lopullista hyväksymistä. Kuntiin pa-lautus jouduttiin usein suorittamaan kahdesti, harvoin kolmeen kertaan.

Kun T-lomakkeet olivat käyneet läpi koko tarkistusmenettelyn ja tulleet siinä hyväksytyiksi, suoritettiin niiden laskenta ja lopullinen kodi-fiointi. Lomakkeen ajankäyttötaulukon kuukausimerkinnät laskettiin vaakasuoraan yhteen, saadut summat laskettiin edelleen lomakkeen osoit-tamalla tavalla ryhmitellen yhteen, ja näin saatujen summien koodit merkittiin niille varattuihin paikkoihin lomakkeen oikeassa reunassa. Tässä kodifioinnissa käytettiin seuraavaa merkintää (luokitusta):

Arkipäiviä	Koodi
0	00
1—10	01
11—20	02
21—30	03

jne.

Koska mainittu työ oli suoritettava pääsälaskuna (tosin tarkoitukseen kehitettyä šablonia apuna käyttäen), sen virhemahdollisuudet olivat melkoiset. Virheiden karsimiseksi metsänhoitaja T. S. NYMALM tarkisti koko käsitellystä aineistosta joka kymmenennen lomakkeen. Jos virhemer-kintöjä 100:ssä lomakkeessa oli enemmän kuin yksi, tarkistettiin asian-omaisen työntekijän suorittama työ kokonaisuudessaan, ja mikäli yhtä suuri virhesadannes toistui, hänet erotettiin.

3222. Reikäkorttiaineistot ja niiden käsittely

Norrbotenin maatalouslaskennassa työvoima-aineiston tulokset las-kettiin pääasiassa käsin, vaikka tutkimuksen muissa osissa käytettiin tilastokoneita (WIESLANDER 1948, s. 6). Tästä syystä lienee tutkimuksessa saatu käytetyksi vain vähäinen osa tällaisen aineiston suomista mahdolli-suuksista. Siitä huolimatta käsittely vei pitkän ajan ja lienee tullut odo-tettua kalliimmaksi. Näin ollen maaseudun työvoiman tutkimuksessa alun perin asetettiin päämääräksi Hollerith-(IBM-) menetelmän soveltaminen aineiston käsittelyyn.

Perusaineistosta (jolla seuraavassa tarkoitetaan T-lomakkeita) laadi-ttiin kaksi toisistaan olennaisesti poikkeavaa reikäkorttiaineistoa. *Henkilö-korttiaineisto* (yhteensä 46 853 korttia) tehtiin lävistämällä monistavin lävistyskonein kunkin T-lomakkeen 1. ja 3. sivun oikeaan reunaan mer-kityt koodit yhdelle reikäkortille. Lävistyksien tarkistus suoritettiin asian-

omaisin konein. Henkilökorttiaineistoa, joka suhteellisen pienen kokonsa vuoksi oli nopeasti lajiteltavissa laskevalla lajittelijalla, käytettiin kaikissa niissä tapauksissa, jolloin koehenkilöä laskentayksikkönä käyttäen tutkittiin näytteen jakautumista ryhmiin koehenkilöiden asuinpaikan, työsuhteen, työkyvyn tai vuoden metsä- ym. työpanoksen suuruuden perusteella.

Henkilökortin lisäksi lävistettiin jokaisesta T-lomakkeesta yksi tai useampia *aikakortteja* mark sensing-menetelmää käyttäen seuraavasti. Erityisille tarkoitusta varten painetuille mark sensing-reikäkortteille siirrettiin kuukausittaiset ajankäyttömerkinnät sekä ao. ajankäyttöryhmän koodi kultakin sellaiselta lomakkeen riviltä (lomakkeen sivu 3), johon oli tehty hyväksytyjä merkintöjä. Merkinnät tehtiin käsin lyijykynällä ja niiden tarkistus suoritettiin 10 %:n pistokokein samalla tavoin kuin lomakkeiden laskennan tarkistus.

Liitteisiin jäljennetyn T-lomakkeen henkilökortista ja yhdestä aikakortista on liitteissä myös valokuvat.

Aikakortit lävistettiin tämän jälkeen mark sensing-koneella siten, että henkilökorttia matriisina käyttäen kunkin lomakkeen aikakortteihin tuli lävistetyksi ao. henkilökortin tiedot sekä lisäksi ajankäyttöryhmän numero (j) ja sen päivien määrät (x_{ij}) kuukausittain. Koneessa käytettiin kaksoislävistykset ja tyhjät sarakkeet estävää kontrollikytkentää.

Täten saatua *aikakorttiaineistoa* (n. 300 000 reikäkorttia) käytettiin niihin tutkimuksiin, joissa päivää (oikeammin 1 000 päivää) laskentayksikkönä käyttäen haluttiin laskea perusjoukon tai sen osan tunnusluku x'_j tai \bar{p}_j . Koska aikakorttiaineistoon lävistettiin kaikki henkilökortteihin tehdyt merkinnät, voitiin se jakaa samanlaisiin osaryhmiin kuin henkilökorttiaineisto. Koehenkilöiden ilmoittaman asuinkunnan perusteella kortistot voitiin jakaa maantieteellisiin alaryhmiin. Eräänä jakoperusteena käytettiin läänijakoa, koska se oli välttämätön vertailukelpoisuuden saavuttamiseksi virallisiin tilastoihin. Kulkulaitosten ja yleisten töiden ministeriön toivomuksesta päätulokset oli sitä paitsi käsiteltävä työvoimapiireittäin ja vuoden 1949 metsätalouskomiteaa varten useissa tapauksissa metsänhoitolautakuntien toiminta-alueittain. Myös päävesistöalueiden tunnuksat merkittiin varalta reikäkortistoon, vaikka tätä jakoperustetta ei olekaan tätä kirjoitettaessa tarvinnut käyttää.

Työkykyisten miesten aikakorttiaineisto siirrettiin, laskemalla kunnittain kaikki saman ajankäyttöryhmän (j) kuukausittaiset x_{ij} -luvut yhteen, summakortteille, niissä olevat luvut x_j kerrottiin suurennustekijällä k (ks. luku 323), kuukausien luvut x'_j laskettiin yhteen, ja lopuksi kaikki

näin saadut kuukausien ja koko vuoden x'_j -luvut taulukoitiin alueittain ja koko valtakunnasta. Samoin meneteltiin tarpeen mukaan aineiston muissa alaryhmissä. Kaikissa mainituissa käsittelyissä käytettiin reikäkorttikoneita, ja kontrollikytkennät suoritettiin mahdollisuuksien mukaan.

323. Suurennusmenetelmä

3231. Suurennus perusjoukon suuruutta tuntematta

Kuten luvussa 3211 mainittiin, valittiin näytteen poimintamenetelmä mm. sitä seikkaa silmällä pitäen, että tutkimuksen päätulokset voitaisiin laskea ennen kuin väestönlaskennasta saatiin tiedot perusjoukon suuruudesta. Yrityksistä huolimatta tätä suunnitelmaa ei voitu tyydyttävästi toteuttaa, koska ensimmäiset väestönlaskennan tulokset 15—64-vuotiaiden miesten lukumääristä osoittivat suunnitelman mukaisesti lasketuissa suurennustekijöissä olevan huomattavia virheellisyyksiä. Lopulliset suurennustekijät oli siten kaikesta huolimatta määritettävä väestönlaskennan tuloksista saadun perusjoukon ja saadun näytteen suhteina.

Kuten seuraavassa osoitetaan, johtui mainittu alkuperäisen suunnitelman epäonnistuminen ennakolta laskemattomasta syntymäpäivien kasautumisilmiöstä. Suomen väestössä on näet teoreettisesti odotettavissa olevaa määrää huomattavasti enemmän sellaisia henkilöitä, jotka merkitsevät tai joille muut ovat virallisiin asiakirjoihin merkinneet syntymäpäiväksi 1. tai 15. päivän.

Koska alkuperäisellä, virheelliseksi havaitulla suurennusmenetelmälläkin on teoreettinen mielenkiintonsa ja lopulliset suurennustekijät laskettiin kustannusten säästämiseksi virheellisiä suurennustekijöitä korjaten, selostetaan seuraavassa sekä alkuperäisten että korjattujen suurennustekijöiden laskutapa.

Selvyyden vuoksi rajoitetaan koko maaseudun työkykyisten miesten kokonaistyöpanosta ja sen rakennetta osoittavien lukujen (HEIKINHEIMO 1954 b) suurennusmenetelmään. Muissa tutkimuksen tuloksissa käytettiin jossain määrin yksinkertaistettua suurennusta.

Vain harvoissa kunnissa T-lomakkeet täytettiin kaikista ohjeissa määritellyistä miehistä. Näin kävi siitä huolimatta, että väestönlaskennan kokemuselostuksen mukaan maalaiskuntien laskentatoimikunnista vain 45 ilmoitti T-miesten löytämisen olleen vaikeata, 175 kuntaa ilmoitti sen olleen helppoa, 264 kunnasta tosin ei saatu lainkaan ilmoitusta tästä sei-

kasta. Laskijoille jokaisen T-lomakkeen täyttämistä maksettua 75 mk:n palkkiota valtaosa toimikunnista piti riittävänä (390 kuntaa). Vain 40 toimikuntaa piti sitä liian pienenä ja 54 pidättyi vastaamasta.

Eräiden arvalla valittujen kuntien henkikirjoista vuodelta 1950 (1. 1. 1950), jotka olivat suurennukseen ryhtyvässä käytettävissä, laskettiin T-miesten lukumäärä kunnassa 1. 1. 1950, siis vuotta ennen väestönlaskennan ja samalla maaseudun työvoiman tutkimuksen teoreettista keräyshetkeä. Näin saatu luku merkitään seuraavassa T_{50} :llä. On huomattava, että T_{50} tarkoittaa vv. 1886—1935 syntyneitä eikä henkikirjoi-tushetkellä 15—64-vuotiaita miehiä. Eri kuntien T_{50} -lukujen vertailu samoista kunnista saatujen T-lomakkeiden lukumääriin osoitti tämän suhteen vaihtelevan niin suuresti, ettei se voinut johtua muusta kuin vastausten kadon osuuden verrattain laajasta vaihtelusta (vastausten katoon palataan luvussa 3311). Tämän takia suurennus päätettiin suorittaa kunnittain, mikä oli välttämätöntä myös sen tähden, että tarvittiin suurennustekijät paitsi lääneittäin myös metsänhoitolautakuntien toiminta-alueittain ja työvoimapiireittäin.

Koska teoreettinen suurennustekijä oli $365.25/24$, oli todellisen suurennustekijän arvioluku k laskettavissa lausekkeesta

$$k = \frac{n_t}{T_h + T_a} \cdot \frac{365.25}{24}, \text{ jossa} \quad (3)$$

n_t on täydellisen näytteen koehenkilöiden lukumäärä, T_h hyväksytyjen T-lomakkeiden lukumäärä näytteessä ja T_a niiden T-lomakkeiden lukumäärä näytteessä, jotka oli hylätty, koska koehenkilö oli puolustuslaitoksen tai rajavartioston vakinaisessa palveluksessa.

Täydellisen näytteen suuruus oli siten riittävän luotettavalla tavalla arvioitava. Se päätettiin suorittaa edellä mainittujen henkikirjaluetteloiden perusteella. Paras tapa olisi luonnollisesti ollut käyttää vuoden 1951 henkikirjoja eli luetteleja kuntien asukkaista 1. 1. 1951. Ne eivät kuitenkaan olleet tarpeeksi ajoissa saatavissa, koska ne valmistuivat vasta syksyllä 1951 ja niitä sitä paitsi tarvittiin verotusluetteleita laadittaessa vuosien 1951 ja 1952 vaihteessa. Tämän takia päätettiin tyytyä vuoden 1950 henkikirjoihin, joiden otteet tilattiin tarkoitusta varten Helsingin yliopiston metsätaloustieteelliseen laitokseen. Niistä laskettiin kunnittain luku T_{50} . Sen perusteella arvioitiin luku T_{51} (T-miesten lukumäärä kunnan henkikirjassa vuoden 1951 alussa). Tämä arvioluku, jota merkitään A_{51} :llä, saatiin seuraavalla päättelyllä.

Vuoden aikana T-miesten lukumäärässä tapahtuneet muutokset ovat *muuttoliikkeen* ja *kuolleisuuden* aiheuttamat. Muuttoliikkeen aiheuttama muutos oli arvioitavissa vain vertaamalla kunnan miespuolista väkilukua 1. 1. 1951 (M_{51}) vastaavaan väkilukuun 1. 1. 1950 (M_{50}). Nämä molemmat luvut olivat riittävän ajoissa saatavissa Tilastollisesta päätoimistosta, vaikka vuoden 1951 henkikirjat eivät olleet yliopiston käytettävissä. Kun näiden lukujen suhteesta M_{51}/M_{50} vähennetään syntyneiden ja kuolleiden miespuolisten henkilöiden erotus v. 1950 jaettuna vuoden 1950 keskiväkiluvulla, p_{se} , saadaan lauseke $M_{51}/M_{50} - p_{se}$, joka likimäärin ilmaisee suhteen T_{51}/T_{50} , kun T-miesten kuolleisuutta vuoden 1950 aikana ei oteta huomioon. Miesten kuolleisuus ikäluokissa 15—64 vuotta laskettuna vuoden 1950 keskiväkiluvusta, merkitään p_{ku} :lla. Lauseke

$$A_{51} = (M_{51}/M_{50} - p_{se}) T_{50} - p_{ku} T_{50}$$

ilmaisee siten likimäärin sen luvun, joka olisi saatu laskemalla T_{51} vuoden 1951 henkikirjoista.

Luku A_{51} ei vielä ilmaise täydellisen näytteen arvioitua suuruutta t_{51} , koska A_{51} sisältää myös laitoshoidokit, joita ei ohjeiden mukaan pitänyt ottaa näytteeseen. Laitoshoidokkien lukumäärän suhde koko 15—64-vuotiaitten miesten lukumäärään 31. 12. 1950 merkitään p_{la} :llä. Siten

$$t_{51} = A_{51} - p_{la} A_{51} \text{ eli}$$

$$t_{51} = (M_{51}/M_{50} - p_{se}) T_{50} - p_{ku} T_{50} - p_{la} [(M_{51}/M_{50} - p_{se}) T_{50} - p_{ku} T_{50}]$$

Laskennan helpottamiseksi annettiin kaavalle likiarvomuoto

$$t_{51} = [M_{51}/M_{50} - (p_{se} + p_{ku} + p_{la})] T_{50} \quad (4)$$

Arvioluku t_{51} laskettiin kunnittain. Luku p_{se} oli saatavissa Tilastollisesta päätoimistosta vain lääneittäin siten, että se Uudenmaan läänissä tarkoitti maalaiskuntien, muissa lääneissä maalaiskuntien ja kauppaloiden miespuolista väestöä. Luku p_{ku} saatiin samasta lähteestä maalaiskuntien ja kauppaloiden miespuoliselle 15—64-vuotiaalle väestölle vain keskimäärin koko maassa, se oli 0.0076. Myös p_{la} saatiin vain koko maaseudun keskiarvona. Laitoshoidokkien määrän laskelmat perustettiin saatavissa oleviin tietoihin kunnalliskodeissa, alkoholistiparantoloissa, irtolaishuoltolaitoksissa (Sosiaalisen tutkimustoimiston antamia lukuja), tuberkuloosiparantoloissa (parantoloiden vuosikertomukset), mielisairaaloissa (mieli-

sairaaloiden vuosikertomukset), sairaaloissa (Suomen tilastollinen vuosikirja 1953, s. 58 ja lääketieteen kandidaatti K. JÄNTIN antamia tietoja) sekä vankiloissa (oikeusrekisterin ja vankeinhoitotoimiston tietoja) lähimain vv:n 1950—51 vaihteessa olleista 15—64-vuotiaista maalaiskunnissa kirjoille merkityistä miehistä. Laskelmien lopputulos oli 12 092 miestä. Jakamalla tämä luku väestönlaskennan ennakkotietojen perusteella arvioidulla maalaiskuntien kaikkien 15—64-vuotiaiden miesten lukumäärällä saatiin $p_{1a} = 0.0143$.

Kaavan 4/s. 69 johdossa voidaan luetella lukuisia epätasällisyyksiä, jotka kuitenkin eivät olleet vältettävissä. Niinpä siinä edellytetään, että kunnan muuttovoitto tai -tappio ikäluokissa 15—64-vuotta oli suhteellisesti sama kuin koko miespuolisella väestöllä, mikä harvoin pitänee paikkaansa. Koko läänin maalais- ja kauppalakuntien syntyvyyden enemmisyssuhdeluvun käyttäminen kunnittain oli myös epätasällistä. Vielä epätarkempaa oli käyttää koko valtakunnan (maalais- ja kauppalakuntien) kuolleisuussuhdelukua ja laitoshoidokkien suhdelukua kunnittain. Tosin nämä suhdeluvut ovat verrattain pieniä, joten niiden virheet eivät sanottavasti vaikuta yhden kokonaisen tarkkuudella laskettuihin t_{51} -lukuihin.

Kunnan todellisen suurennustekijän arvio k saatiin tämän jälkeen asettamalla kaavassa 3/s. 68 n_t -tekijän sijaan t_{51} .

Tulokset laskettiin tämän jälkeen luvussa 3222 mainitulla tavalla suorittamalla reikäkorttikoneissa suurennus kuntien aikakorteissa ja laske-malla tabulaattorilla tulot yhteen kulloinkin tarvittaville kuntaryhmille sekä koko valtakunnalle.

3232. Lopullinen suurennusmenetelmä

Edellä selostetun laskentamenettelyn jälkeen ryhdyttiin varsinaiseen tulosten käsittelyyn, jonka aikana väestönlaskennan ikäluokkataulukot alkoivat valmistua. Täten tuli mahdolliseksi verrata väestönlaskennan tietojen perusteella arvioitua perusjoukon suuruutta näytteestä suurentamalla saatuun perusjoukon henkilölukuun.

Väestönlaskennan ikäluokkatiedoista poimittiin lääneittäin maalaiskuntien (ja Karhulan kauppalan) 15—64-vuotiaan miespuolisen väestön suuruutta osoittavat luvut. Näistä vähennettiin p_{1a} % (jonka korjattu arvo oli 1.47 %) eli arvioidun laitoshoidossa olleen väestön osuus maalaiskuntien koko 15—64-vuotiaasta miespuolisesta väestöstä. Näin saadut luvut osoittivat perusjoukon todellisen suuruuden N niin tarkasti, kuin

Taulukko 9. Estimoidun perusjoukon (n') ja todellisen perusjoukon (N) suuruuden vertailu. Päätutkimus.

Table 9. Comparison of estimated population (n') and actual population (N). Main Survey.

Lääni County	n' miestä males	N miestä males	$\frac{n' - N}{N}$ %
Uudenmaan — <i>Uusimaa</i>	72 053	67 268	+ 7
Turun ja Porin — <i>Turku and Pori</i>	135 557	131 590	+ 3
Ahvenanmaa — <i>Aaland</i>	6 125	5 707	+ 7
Hämeen — <i>Häme</i>	100 699	95 128	+ 6
Kymen — <i>Kymi</i>	68 723	65 171	+ 5
Mikkelin — <i>Mikkeli</i>	65 934	59 488	+ 11
Kuopion — <i>Kuopio</i>	138 029	119 309	+ 16
Vaasan — <i>Vaasa</i>	155 378	144 384	+ 8
Oulun — <i>Oulu</i>	99 234	89 819	+ 10
Lapin — <i>Lapland</i>	42 158	37 429	+ 13
Koko maaseutu — <i>All rural districts</i>	883 890	815 293	+ 8.4

oli mahdollista se laskea. Saatuja lukuja verrattiin vastaaville alueille koehenkilöiden lukumääristä laskettuihin perusjoukon arviolukuihin $n' = k(T_h + T_a)$.

Tulokset esitetään taulukossa 9. Näytteestä suurentamalla saatu perusjoukko n' oli kaikissa lääneissä huomattavasti suurempi kuin N . Koko maassa keskimäärin virhe oli 8 % N :stä ja maan itä- ja pohjoisosissa se oli yli 10 %. Vastaavansuuruisen ja -suuntaisen virheen täytyi siis olla myös kaikissa kertoimilla k suurennetuissa absoluuttisissa luvuissa.

Näin suuria virheitä ei tietenkään voitu jättää korjaamatta. Toisaalta olisi koko suurennusmenetelmän uudistaminen kunnittaisista tuloksista lähtien vienyt niin paljon aikaa ja vaatinut sellaisia kustannuksia, että sitä oli mahdoton toteuttaa. Jo valmiiksi lasketut taulukot päätettiin sen tähden korjata lääneittäin, metsänhoitolautakuntien toiminta-alueittain ja työvoimapiireittäin lasketuilla kertoimilla N/n' .

Lopulliset tulokset tulivat täten suurennetuiksi läänien, metsänhoitolautakuntien toiminta-alueiden ja työvoimapiirien todellisilla suurennuskertoimilla $K = N/(T_h + T_a)$. Kertoimet K ja todelliset otantasuhteet $F = \frac{1}{K}$ on merkitty lääneittäin taulukkoon 10/s. 72.

Taulukko 10. Lopulliset suurennuskertoimet (K) ja otantaosuudet ($F = 1/K$). Pää tutkimus.

Table 10. Final raising factors (K) and sampling ratios ($F = 1/K$). Main Survey

Lääni — County	K	F %
Uudenmaan — Uusimaa	18.8584	5.30
Turun ja Porin — Turku and Pori	17.3762	5.75
Ahvenanmaa — Åland	17.8903	5.59
Hämeen — Häme	17.8544	5.60
Kymen — Kymi	17.7336	5.64
Mikkelin — Mikkelä	16.6914	5.99
Kuopion — Kuopio	16.5960	6.03
Vaasan — Vaasa	17.3768	5.75
Oulun — Oulu	16.7448	5.97
Lapin — Lapland	17.4902	5.72
Koko maaseutu		
All rural districts	17.3363	5.77

Tarkistuksen vuoksi laskettiin lääneittäin luvut z' myös siten, että korjaus suoritettiin kunnittain. Tulos osoitti, että korjausmenetelmä ainakin mainittujen lukujen osalta johti niin pieniin virheisiin, etteivät ne näy taulukkojen katkaistuissa luvuissa.

3233. Syntymäpäivämerkintöjen kasautumisilmiö

Ennen suurennustekijöiden k laskentaa tarkistettiin, johtaako kaava 4/s. 69 todella käyttökelpoiseen, vuoden 1951 (1. 1. 1951) henkikirjoista saatujen T-miesten lukumäärän likiarvoon. Tätä varten suoritettiin Lohjan, Tuusulan ja Helsingin kihlakunnissa (yhteensä 20 kuntaa) rinnan vuoden 1950 henkikirjojen T-miesten laskennan kanssa myös sama laskenta vuoden 1951 henkikirjoista, jotka olivat luonnosasteisina aivan lyhyeksi ajaksi saatavissa ko. tarkoitukseen. Tällöin todettiin vuoden 1951 henkikirjoista laskettujen T-miesten lukumäärien poikkeavan vuoden 1950 henkikirjojen perusteella lasketuista likiarvoista (laitoshoidokkeja ei tällöin niistä vähennetty) eri kunnissa + 21 %:sta — 11 %:iin. Mainittujen kolmen kihlakunnan alueella yhteensä oli T_{51} 2.3% arviolukua A_{51} suurempi. Näin suppean aineiston perusteella ($T_{51} = 3323$) saatua tulosta ei tosin voida yleistää, mutta se näyttää viittaavan siihen, että kaava 4/s. 69 mah-

dollisesti antaa keskimäärin liian suuria tuloksia. Syyksi voisi lähinnä olettaa, että kunnista ja kuntiin muuttaneissa olisi enemmän 15—64-vuotiaita kuin koko miespuolisessa väestössä. Tällöin tulisi sanotun virheen olla positiivinen silloin, kun kunnasta on muuttanut enemmän miespuolista väestöä kuin sinne on tullut muista kunnista ja päinvastaisessa tapauksessa negatiivinen. Tarkistetuissa kunnissa ei kuitenkaan havaittu olevan korrelaatiota näiden seikkojen välillä. Sen tähden oletettiin suoritettuna kokeilussa ilmenneen eron johtuneen satunnaisvaihtelusta.

Sen seikan selvittämiseksi, onko kuukauden 1. tai 15. päivän syntymäpäiväkseen ilmoittaneiden henkilöiden suhteellinen osuus jostakin suuresta henkilöryhmästä meidän maassamme odotettavissa oleva 24/365.25, suostui tekijän pyynnöstä Kansaneläkelaitoksen vakuutus teknillinen osasto muiden töittensä yhteydessä laskemaan tämän suhteen osasta vakuutuskortistiaan (laskennan suoritti valtiotieteiden maisteri ERKKI JÄRVINEN). Yhteensä 407 767 tutkitun eri puolilla maata valituista kunnista kotoisin olevan vv. 1890—1935 syntyneen mies- ja naispuolisen vakuutetun joukossa todettiin olevan 28 631 henkilöä, joiden syntymäpäiväksi oli merkitty jonkin kuukauden 1. tai 15. päivä. Saatu luku ylittää odotusarvon 6.87 ± 0.19 %:lla, joten erotus on erittäin merkitsevä. Osasta Helsingin kaupunkia ja Nokian kauppalasta kotoisin olevien eläkkeenmaksajien joukossa mainittuina päivinä syntyneitä oli jokseenkin täsmälleen odotusarvoa vastaava määrä. Maalaiskunnissa ero oli siten suurempi kuin asutuskeskuksissa.

Kansaneläkelaitoksen saamat tiedot eläkemaksun suorittajien syntymäajoista perustuvat henkikirjatietoihin. Tosin laitoksen asiamiesten velvollisuuksiin kuuluu tarkistaa tiedot, mutta tämä tuskin tapahtuu kovin usein vähämerkityksisinä pidettäviin syntymäpäiviin nähden. Näin ollen näyttää selvältä, että maassamme henkikirjoihin on merkitty tässä kyseeseen tulevissa ikäluokissa kuukauden 1. ja 15. päivänä syntyneiksi huomattavasti suurempi joukko henkilöitä kuin teoreettisesti olisi odotettavissa. Edellä selostettu näytteestä suurentamalla saadun perusjoukon ja väestönlaskennan tietojen perusteella arvioidun perusjoukon ero johtuu ilmeisesti tästä seikasta.

Luvussa 3211 todettiin, ettei ole järjellistä syytä olettamukseen, että synnytyksillä olisi taipumus kasautua tilastollisesti merkitsevällä tavalla kuukauden 1. tai 15. päivän enempää kuin muidenkaan päivien kohdalle. Mainitun ilmiön täytynee siten johtua henkikirjoihin ja muihin asiakirjoihin tehdyistä virheellisistä syntymäpäivämerkinnöistä siten, että virhemerkinnöillä on mainitunlainen kasautumistaipumus.

Useat tutkijat ovat todenneet, että ihmisellä on voimakas taipumus pyöristää lukuja ennen kaikkea 0:lla ja 5:llä päättyviksi. Niinpä HIESS (1931, ss. 51—64) ja BACHI (1952) ovat tutkineet ilmiötä eri maiden väestönlaskentojen ikäjakautumisissa ja todenneet, että kun ikä on ilmoitettava vuosissa, numeroon 0 ja 5 päättyvät luvut ovat suosittuja miltei kaikissa tutkituissa maissa. Myös numeroon 2, 3 ja 8 päättyvillä luvuilla on eräissä maissa lievä kasautumistaipumus. Sen sijaan numeroon 1, 4, 6, 7 ja 9 päättyviä lukuja kartetaan. Kasautuminen on erittäin voimakas esim. Lähi-Idän maissa, mutta selvästi havaittava myös Australiassa, Englannissa ja USA:ssa. Naisilla se on yleensä selvempi kuin miehillä ja vanhemmilla ikäluokilla voimakkaampi kuin nuorilla. Se on sitä suurempi, kuta enemmän on lukutaidottomia (täysi-ikäisistä) ja kuta vähemmän on tapana viettää syntymäpäivää. BACHI ehdottaa, että tämä ilmiö otettaisiin huomioon erityisin korjausmenetelmin väestönlaskennoissa.

Myöskin tämän tutkimuksen näytteen poiminnassa suoritettu tunnuslukujen valinta (1 ja 15), johon asiantuntijoillakaan ei ollut huomauttamista, on todisteena samasta ilmiöstä.

Tällaisella kasautumisilmiöllä on useita mahdollisuuksia päästä vaikuttamaan nimenomaan henkikirjojen syntymäaikatietoihin. Henkikirjoituslomakkeen täyttäneen useimmiten perheenisä, jonka kyky muistaa perheensä jäsenten syntymäajat, etenkin päivämäärät, on tunnetusti varsin heikko, joten virhemerkintöjä tulee helposti. Henkikirjoittajien toimistoissa samaa perusluetteloa korjataan kolmen vuoden ajan, jonka jälkeen korjattu luettelo kirjoitetaan puhtaaksi. Joka kolmas vuosi on jatkuvastikin samassa kunnassa asuvan henkilön syntymäajalla mahdollisuus jäljennös-
virheen takia muuttua. Tämän lisäksi on todennäköistä, ettei varsinkaan iäkkäiden henkilöiden tarkka syntymäpäivä ole aina enempää heidän itsensä muistissa kuin kirkonkirjoissakaan täsmällisesti oikea. Kun kaikilla näillä virhemerkinnöillä on taipumus kasautua juuri määrättyjen lukujen kohdalle, on ymmärrettävissä, että ne johtavat edellä selostettuun ilmiöön.

Syntymäpäivää satunnaislukuna käyttävä systemaattinen poimintamenetelmä näyttää erilaisissa väestöä koskevissa tutkimuksissa ainakin Suomessa käytännöllisyytensä vuoksi yleistyvän maaseudun työvoiman tutkimuksen tultua tunnetuksi. Niinpä muuttotilastossamme käytetään syntymäpäiviä 5, 15 ja 25 (FOUGSTEDT 1953, s. 356) ja Helsingin kaupungin eräässä nuorisotutkimuksessa käytettiin lukua 15 (BRUUN 1952, s. 197) ja vanhustutkimuksessa arvottuja lukuja 5, 13, 16 ja 27 (BRUUN 1953, s. 74). Tämän takia on tärkeää valita kyseiseen tarkoitukseen sellai-

set luvut, jotka antavat mahdollisimman vähän virheellisiä tuloksia. Tosin »vaarallisten» lukujen käyttäminen tuottaa eniten haittaa sellaisessa tapauksessa, jossa suurennustekijät lasketaan perusjoukon suuruutta tuntematta. Kuitenkin on mahdollista ja todennäköistä, että mainittu kasautumisilmiö on erisuuruinen eri yhteiskuntaluokkien ja ikäryhmien syntymäaikamerkinnöissä. Täten niiden käyttäminen voi aiheuttaa itse näytteen-
saamisen vaikeasti todettavia ja korjattavia harhoja.

Päätutkimuksen aineistossa yritettiin Uudenmaan ja Lapin läänien T-lomakkeista tutkia, onko syntymäpäivien 1 ja 15 välillä puheena olevassa suhteessa eroa. Tulos oli seuraava:

	Yhteensä koehenkilöitä	Näistä 1. päivänä syntyneitä koehenkilöitä	%
Uudenmaan lääni	3 431	1 822	53.1 ± 0.8 %
Lapin lääni	1 975	1 040	52.7 ± 1.5 %

Molemmissa tapauksissa oli 1. päivänä syntyneiksi merkittyjen koehenkilöiden määrä huomattavasti suurempi kuin 15. päivänä syntyneiksi ilmoitettujen koehenkilöiden. Testaus osoitti suhteellisen osuuden Uudenmaan läänissä merkitsevästi poikkeavan odotusarvosta 50 %.

Saman seikan selvittämiseksi laskettiin kansaneläkelaitoksen kortistosta 332 877 vakuritetun 1. ja 15. päivinä syntyneiksi merkittyjen lukumäärät, kumpikin erikseen. Todettiin edellisiä olevan 11 998 henkilöä, mikä 9.72 ± 0.97 %:lla ylittää odotusarvon ja jälkimmäisiä 11 292 henkilöä, mikä ylittää odotusarvon 3.26 ± 1.93 %:lla. Näin ollen tämäkin tutkimus hyvin vakuuttavasti osoitti kasautumisilmiön olevan huomattavasti suurempi 1. päivänä syntyneiksi merkityissä. Sadanneslukujen 9.72 ja 3.26 erotuksen 6.46 % keskivirhe on ± 1.09 sadannesyksikköä, joten erotus on erittäin merkitsevä. Näennäisesti tämä tulos poikkeaa BACHIN toteamuksesta, joiden mukaan numeroon 1 päättyviä lukuja kartetaan. Numeroon 1 päättyvien lukujen joukossa on 1 kuitenkin erikoisasemassa ja se on syntymäpäivänä helposti muistettava, lähinnä 0:lla päättyvään lukuun verrattava.

33. Päättökimuksen tulosten luotettavuus

331. Poimintaharhat

3311. Vastausten kadon aiheuttamat harhat

DALENIUS (1949, ss. 13—29) jakaa poimintatutkimuksen (ja periaatteessa myös täydellisen tutkimuksen) virhemahdollisuudet kolmeen pääryhmään: poiminnasta, mittauksesta ja käsittelystä johtuvat virheet. Tulosten luotettavuuden kannalta on vielä tärkeämpi jakaa virheet satunnaisvirheisiin, jotka näytteen suuretessa pyrkivät yhä paremmin tasoittamaan toisensa, ja harhoihin (systemaattiset virheet). Seuraavassa käsitellään sen tähden kaikki satunnaisvaihtelut yhdessä.

Väestöstä poimitun näytteen kaikista yksiköistä saadaan harvoin vaadittavat tiedot. Vastausten puuttuminen (kato) muodostaakin sen tähden otantatutkimuksen erään vaikeimmin ratkaistavan ongelman. Kuten luvussa 2431 mainittiin, saatiin maaseudun työvoiman tutkimuksessa vain harvoista kunnista täydellinen näyte. Vastausten kadon osuus (J_{max}) mitattuna luvun t_{51} ja summan $T_h + T_a$ erotuksen suhteena t_{51} :een

$$J_{max} = 1 - (T_h + T_a)/t_{51}$$

vaihteli suuren Espoon kunnan 61.7 %:sta 0 %:iin (tai laskentamenetelmästä johtuen sitä pienempään lukuun) 31:ssä kunnassa. Lääneittäin se oli taulukossa 11 esitetyn suuruinen.

Sanottu laskentatapa johtaa liian suuriin sadanneslukuihin, koska väestönlaskijat saattoivat useissa tapauksissa todeta, ettei henkikirjojen mukaan näytteeseen kuuluvan henkilön syntymäpäivä ollutkaan 1. tai 15. päivä, ja asianomaisesta ei niin muodoin täytetty T-lomaketta.

Kadon osuus voidaan laskea toisellakin tavalla. Kun perusjoukon suuruus tunnetaan, olisi täydellisen näytteen teoriassa pitänyt olla $\frac{24}{365.25} N$.

Kun saadun näytteen $T_h + T_a$ suuruutta verrataan siihen, saadaan kadon osuus

$$J_{min} = 1 - 365.25 (T_h + T_a) / 24 N$$

Taulukkoon 11 on merkitty myöskin viimeksi mainitulla kaavalla lasketut kadon osuudet. Luvut J_{min} ovat huomattavasti pienemmät kuin

Taulukko 11. Estimoidut vastausten kadon osuudet. Päättökimuk.

Table 11. Estimated non-response ratios. Main Survey.

Lääni — County	J_{max} %	J_{min} %
Uudenmaan — Uusimaa	24.7	19.3
Turun ja Porin — Turku and Pori	15.1	12.4
Ahvenanmaa — Åland	20.8	14.9
Hämeen — Häme	19.5	14.8
Kymen — Kymi	18.8	14.2
Mikkelin — Mikkelä	17.6	8.8
Kuopion — Kuopio	20.6	8.3
Vaasan — Vaasa	18.7	12.4
Oulun — Oulu	17.9	9.1
Lapin — Lapland	22.9	13.0
Koko maaseutu — All rural districts	19.1	12.2

J_{max} . Varsinkin Pohjois- ja Itä-Suomen lääneissä ero on suuri. Todellinen kadon osuus ei voi olla merkittävästi suurempi kuin J_{max} , ja kuta suurempi on alueen väestö, sitä vähemmän on todennäköistä, että todellinen kadon osuus on pienempi kuin J_{min} . On vaikea päätellä, kumpi on lähempänä todellista arvoa. Se riippuu siitä, kuinka suuressa määrin näytteeseen sisältyi »vale-T-miehiä», siis sellaisia, jotka todellisuudessa eivät olleetkaan syntyneet 1. tai 15. päivänä.

Joka tapauksessa on tuntemattomien näyteyksiköiden määrä huolestuttavan suuri, tosin ei läheskään niin suuri kuin on tavallista esim. USA:ssa paljon käytetyissä postitse suoritetuissa tiedusteluissa (The Agricultural ... 1949). Koska olettamukseen, että epätäydelliseen näytteeseen saadut koehenkilöt edustavat koko näytettä, sisältyy aina tuntematonta suuruusluokkaa olevan harhan mahdollisuus, hyväksyttävässä otantamenetellessä vaaditaan erikoistutkimus puuttuvien näyteyksiköiden laadusta. Tämä tapahtuu parhaiten ottamalla niistä jollakin yksinkertaisella menetelmällä otos ja tutkimalla se samanlaisella mittaustmenetelmällä kuin alkuperäisenkin näyte (esim. HANSEN ja HURWITZ 1946). Väestönlaskennan yhteydessä mainittua tapaa ei kuitenkaan voitu käyttää, vaan työvoimatutkimuksen puuttuvaa näytteen osaa oli pyrittävä tutkimaan muilla, huomattavasti epävarmemmilla keinoilla.

Muutamissa kunnissa, joissa kadon osuus oli erityisen suuri, yritettiin tosin henkikirjatietojen perusteella selvittää näytteestä puuttuvien koe-

henkilöiden ammattijakautuma. Tulokset eivät osoittaneet huomattavia eroja näytteeseen saatujen ja siitä puuttuvien vuoden 1950 henkikirjoihin merkittyjen T-miesten ammattijakautumassa. Tämä yritys keskeytettiin, koska henkikirjojen ammattinimikkeet olivat siinä määrin epäselvät ja tulkinnanvaraiset, ettei vertailun tuloksille voitu antaa käytännön arvoa.

Kadon aiheuttaman harhan tutkimukset oli sen tähden perustettava saadun näytteen ja väestönlaskennasta saadun vastaavan ikäryhmän ikä- ja ammattijakautuman vertailuun. Tosin mainitut jakautumat eivät nekään olleet täysin vertailukelpoiset, koska väestönlaskennan tiedoista saatiin ainoastaan koko 15—64-vuotiaan miespuolisen väestön ikä- ja sen ammatissa toimivan osan ammattijakautuma alueittain. Niihin sisältyvät myös laitoshoidokit ja ammattisotilaat. Kumpikaan viimeksi mainituista ryhmistä ei sisälly näytteeseen (T_h). Tämän takia väestönlaskennasta saatua ryhmää seuraavassa nimitetään »perusjoukoksi» (Pe) erotukseksi todellisesta perusjoukosta (N).

Ikäjakautuman vertailun tulokset näkyvät taulukosta 12 ja kuvasta 3/s. 80.

Koko maaseudun tuloksia tarkasteltaessa vertailu näyttää osoittavan, että molemmissa tapauksissa *ikäjakautuma on varsin tyydyttävällä tavalla yhtäpitävä* (läänittäiset näytteen luvut painotettiin koko maaseudun sadanneksia laskettaessa ao. läänien lopullisilla suurennustekijöillä). Tosin ikäluokissa alle 25 vuoden on näytteessä hiukan liian vähän ja vastaavasti ikäluokissa yli 50 vuoden liian runsaasti koehenkilöitä. Erot ovat kuitenkin niin pieniä, ettei niillä ole sanottavaa vaikutusta tutkimuksen tuloksiin, varsinkin kun todettiin 15—19-vuotiaiden nuorukaisten ajankäyttö rakenteeltaan hyvin samantapaiseksi kuin 20—64-vuotiaiden miesten ajankäyttö.

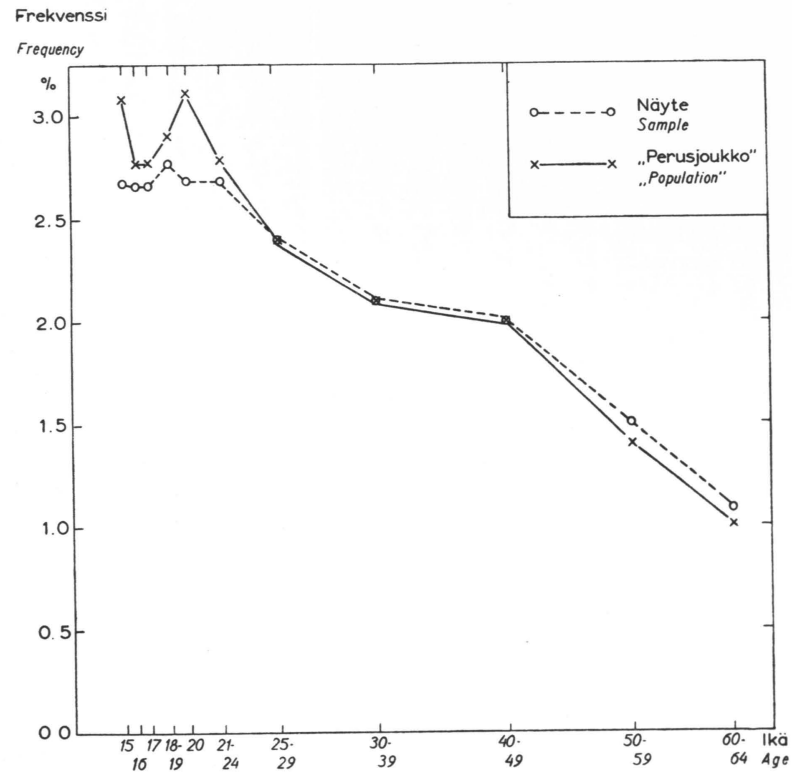
Lääneittäin taulukkoa tarkasteltaessa erot ovat yleensä samansuuntaiset kuin koko valtakunnassa eivätkä Ahvenanmaata lukuun ottamatta myöskään ole kovin suuret. Ahvenanmaan suurehkot erot taas selittyvät koehenkilöiden pienestä luvusta (vain 319 koehenkilöä). Yleensäkin erot eivät ylitä kolminkertaisen keskivirheen rajaa.

Näytteen ja »perusjoukon» ammattijakautumaa yritettiin Uudenmaan ja Lapin lääneissä verrata seuraavalla menettelytavalla. Väestönlaskennassa ammattikoodien merkitsijöinä toimineet kaksi luotettaviksi työntekijöiksi havaittua henkilöä merkitsi Uudenmaan läänin hyväksytyihin T-lomakkeisiin väestönlaskennassa käytetyn ammattinimikkeistön koodin. He yrittivät mahdollisuuksien mukaan T-lomakkeesta päätellä, mikä ammattinimike kyseiselle henkilölle oli merkitty varsinaisen väestönlasken-

Taulukko 12. Miesten ikäjakautuma sadanneksina. »Perusjoukon» (Pe) ja näytteen (n) vertailu. Päätutkimus.

Table 12. Percentage distribution of males into age classes. Comparison between »population» (Pe) and sample (n). Main Survey.

Lääni — County	Ikäluokka — Age class										Yhteensä — Total		
	15	16	17	18—19	20	21—24	25—29	30—39	40—49	50—59	60—64	%	Koehenkilöitä Males in sample
Uudenmaan <i>Uusimaa</i>	2.5 1.9	2.3 1.8	2.3 2.3	4.9 5.0	2.6 2.7	10.3 10.1	11.8 11.6	20.7 19.9	22.0 22.5	15.0 16.6	5.6 5.6	100 100	3 557
Turun ja Poriin <i>Turku and Pori</i>	2.9 2.5	2.8 2.7	2.6 2.6	5.7 5.5	3.1 2.5	10.9 10.7	11.8 11.0	19.6 20.1	20.8 21.4	14.6 15.6	5.2 5.4	100 100	7 546
Ahvenanmaa <i>Aaland</i>	2.8 3.4	2.5 0.6	2.6 2.8	5.5 4.4	2.3 4.4	9.5 10.7	11.7 11.6	22.1 23.2	19.2 16.0	16.5 16.0	5.3 6.9	100 100	319
Hämeen <i>Häme</i>	2.8 2.6	2.6 2.7	2.6 2.7	5.4 5.2	3.0 2.4	10.7 9.6	11.7 12.2	19.8 20.2	21.0 21.8	14.6 15.2	5.8 5.4	100 100	5 302
Kymen <i>Kymi</i>	2.7 2.4	2.5 2.4	2.5 1.9	5.4 6.2	2.8 2.4	10.8 10.6	11.7 11.4	20.3 20.1	20.9 20.9	15.0 16.1	5.4 5.6	100 100	3 645
Mikkelin <i>Mikkeli</i>	3.1 2.5	2.9 2.6	2.9 2.8	5.7 5.5	3.0 2.6	11.0 10.6	11.3 12.3	19.2 20.1	19.9 20.0	15.1 15.4	5.9 5.6	100 100	3 563
Kuopion <i>Kuopio</i>	3.3 2.8	3.2 2.9	2.9 2.3	6.1 5.5	3.3 3.0	11.5 10.7	11.8 12.0	20.0 21.3	19.5 20.1	13.5 14.4	4.9 5.0	100 100	7 158
Vaasan <i>Vaasa</i>	3.2 3.2	3.0 2.7	2.9 2.6	6.0 5.7	3.1 3.1	11.3 10.3	12.4 12.4	21.2 21.4	19.1 20.2	13.2 13.7	4.6 4.7	100 100	8 294
Oulun <i>Oulu</i>	3.5 3.1	3.3 3.0	3.1 3.7	6.3 6.2	3.4 2.8	12.2 11.1	12.9 12.8	20.7 20.2	17.7 18.3	12.0 13.4	4.9 5.4	100 100	5 350
Lapin <i>Lapland</i>	3.4 3.3	3.4 3.8	3.1 3.3	6.3 5.8	3.3 2.2	13.0 12.9	13.4 12.2	20.4 20.5	17.4 17.7	11.8 12.7	4.5 5.6	100 100	2 116
Koko maaseutu <i>All rural districts</i>	3.1 2.7	2.8 2.7	2.8 2.7	5.8 5.6	3.1 2.7	11.1 10.6	12.0 12.0	20.4 20.5	20.1 20.4	13.7 14.8	5.1 5.3	100 100	46 850
	+ 0.4	+ 0.1	+ 0.1	+ 0.2	+ 0.4	+ 0.5	± 0.0	- 0.1	- 0.3	- 1.1	- 0.2	—	—



Kuva 3. Miesten ikäjakautuma sadanneksina. »Perusjoukon» (Pe) ja näytteen (n) vertailu. Päättökäytös. Lähde: taulukko 12/s. 79.

Figure 3. Percentage distribution of males into age classes. Comparison between »population» (Pe) and sample (n). Main survey. Source: Table 12/p. 79.

nan aineistoa käsiteltäessä. Tätä tarkoitusta varten väestönlaskennan laajaa ammattinimikkeistöä kuitenkin yhdistelemällä supistettiin niin, että siihen jäi ainoastaan 10 taulukossa 13 lueteltua ammattiryhmää. Paras tapa olisi ollut hakea koehenkilöiden ruokakuntalomakkeet (C) väestönlaskennan aineistosta ja siirtää ammattikoodi niistä T-lomakkeisiin. Näin ei kuitenkaan voitu tämän menetelmän suurten kustannusten tähden tehdä. Käytettyyn menetelmään jäi tällöin se virhemahdollisuus, että T-lomakkeeseen tuli merkityksi toinen ammattiryhmä, kuin mihin asianomainen henkilö väestönlaskennassa kuului. Tosin annettujen ohjeiden mukaan väestönlaskijoiden oli merkittävä T-lomakkeeseen sama ammatti-

Taulukko 13. Miesten ammattijakautuma sadanneksina. »Perusjoukon» ja näytteen vertailu. Uudenmaan lääni. Päättökäytös.

Table 13. Percentage distribution of males into occupation groups. Comparison between »population» and sample. Uusimaa County. Main Survey.

Ammattiryhmä Occupation group *	»Perus- joukko» »Popu- lation»	Näyte — Sample		Yhteensä Total
		Syntymä- päivä Birthday 1	Syntymä- päivä Birthday 15	
1. Henkisen alan työlliset	12.5	10.7	10.4	10.6
2. Maanviljelijät ja heidän maatilan töihin osallistuvat perheenjäsenensä. Turkkiläisten hoitajat, poronhoitajat ja puutarha-alan työlliset	30.4	32.3	35.7	33.9
3. Maataloustyöntekijät	9.4	9.8	8.9	9.3
4. Metsätalouden ja uiton työntekijät	3.7	6.4	4.7	5.6
5. Metsästäjät ja kalastajat	1.1	1.7	2.1	1.9
6. Kuljetustyölliset	9.2	7.8	8.4	8.1
7. Teollisuuden, käsityön, kaivoksien ja kivilouhimoiden työlliset	26.4	24.7	23.1	23.9
8. Talonrakennustyöntekijät	5.8	4.6	5.0	4.8
9. Katu-, rautatie-, tie-, vesi- ym. rakennusalan työntekijät	0.9	1.0	0.8	0.9
10. Ilmoittamattomat ammatit ja ammatit, joita ei voitu määritellä	0.6	1.0	0.9	1.0
Yhteensä — Total %	100.0	100.0	100.0	100.0
Koehenkilöitä — Males in sample	.	1 821	1 609	3 430

- * 1. White-collar workers.
2. Farmers and their family members working on the farm, fur-farmers, reindeer breeders and gardeners.
3. Paid workers in agriculture.
4. Paid workers in forestry and floating.
5. Hunters and fishers.
6. Transport workers.
7. Workers in industry, handicraft, mining.
8. Paid workers in house construction.
9. Paid workers in street, railway, road, water-way etc. construction.
10. Occupation unstated or undefinable.

nimike, joka koehenkilöiden kohdalla merkittiin C-lomakkeeseenkin (laskettavat saivat vapaasti merkitä ammattinsa harkitsemallaan tavalla). T-lomakkeiden ammattimerkinnät puuttuivat kuitenkin joskus kokonaan ja vielä useammin ne olivat epämääräisiä. Niinpä sekatyömies, työmies ja talollisen poika olivat tavallisia ja vaikeasti tulkittavia nimityksiä. Toisaalta T-lomakkeen ajankäyttötäulukosta juuri mainitunlaisten koe-

henkilöiden pääammatti oli paremmin todettavissa kuin väestönlaskennan C-lomakkeesta, jossa vain metsä- ja uittotöihin käytettyjen kuukausien määrä oli merkitty (toisaalta kyllä C-lomakkeessa päättelyä helpotti työnantajan nimi). Näin ollen maaseutuväestön ammattimerkinnässä oli ollut runsaasti harkinnanvaraisuutta itse väestönlaskennan käsittelyvaiheessa eikä ollut varmaa, että tämä harkinta T-lomakkeiden ammattiryhmiä merkittäessä läheskään aina tapahtui samalla tavalla, kun harkinnan mahdollisuudet pakostakin olivat erilaiset. Lapin läänin T-lomakkeiden merkintä katsottiinkin voitavan yhtä suurella tarkkuudella suorittaa yliopiston metsätaloustieteellisessä laitoksessa väestönlaskennan ammatinmerkintä-ohjeita mahdollisimman tarkoin seuraten.

T-lomakkeet laskettiin tämän jälkeen ammattiryhmittäin ja saatua jakautumaa verrattiin maalaiskuntien 15—64-vuotiaan miespuolisen ammatissa toimineen väestön vastaavaan jakautumaan. Sitä ennen näytteestä vähennettiin sellaiset henkilöt, jotka oli merkitty »ei ammatissa toimivaksi» ja »perusjoukosta» vastaavasti »sotilashenkilöt». Vertailun tulos näkyy taulukoista 13/s. 81 ja 14 sekä kuvasta 4/s. 84.

Uudenmaan läänin vertailu osoittaa, että henkisen alan työllisten ja teollisuustyöntekijöiden ryhmät ovat näytteessä liian pienet ja maanviljelijöiden, metsätyömiesten sekä niihin liittyvien ammattien ryhmät liian suuret. Testi osoittaa, että mainituissa suhteissa näyte merkitsevästi eroaa perusjoukosta. Uudenmaan läänin tutkimustuloksia käytettäessä on mainittu seikka otettava huomioon. Virheen voidaan katsoa johtuneen poikkeuksellisen suuresta kadon osuudesta Helsingin ympäristökunnissa (Espoo ja Helsingin maalaiskunta), joissa asiaa tutkittaessa kävi ilmi, että jotkut laskijat olivat katsoneet työvoimatutkimuksen koskevan etupäässä maa- ja metsätaloustaloustaloutta. Tämän harhan korjauslaskelmia ei voitu suorittaa. Ainoa keino olisi ollut suorittaa mainituissa kunnissa aineiston täydennys. Kuntien laskentatoimikuntien virheellisten ilmoitusten johdosta oltiin kuitenkin kadosta tietämättömiä niin kauan, ettei aineiston täydennyskeräys ollut enää mahdollista.

Lapin läänin vertailu ja sen tulosten testaus ei osoittanut näytteen merkittävästi eroavan perusjoukosta, vaikka ko. läänissä kadon osuus oli lähes samaa suuruusluokkaa kuin Uudenmaan läänissä. Muissa lääneissä, joissa vastausten kato oli vähäisempi, vertailua ei suoritettu. Syynä oli myös se, että ammattinimikkeiden määrittämistä vaikeuksien vuoksi vertailun tulosta ei voitu pitää täysin luotettavana.

Taulukko 14. Miesten ammattijakautuma sadanneksina. »Perusjoukon» ja näytteen vertailu. Lapin lääni. Päättökäsim.

Table 14. Percentage distribution of males into occupation groups. Comparison between »population» and sample. Lapland County. Main Survey.

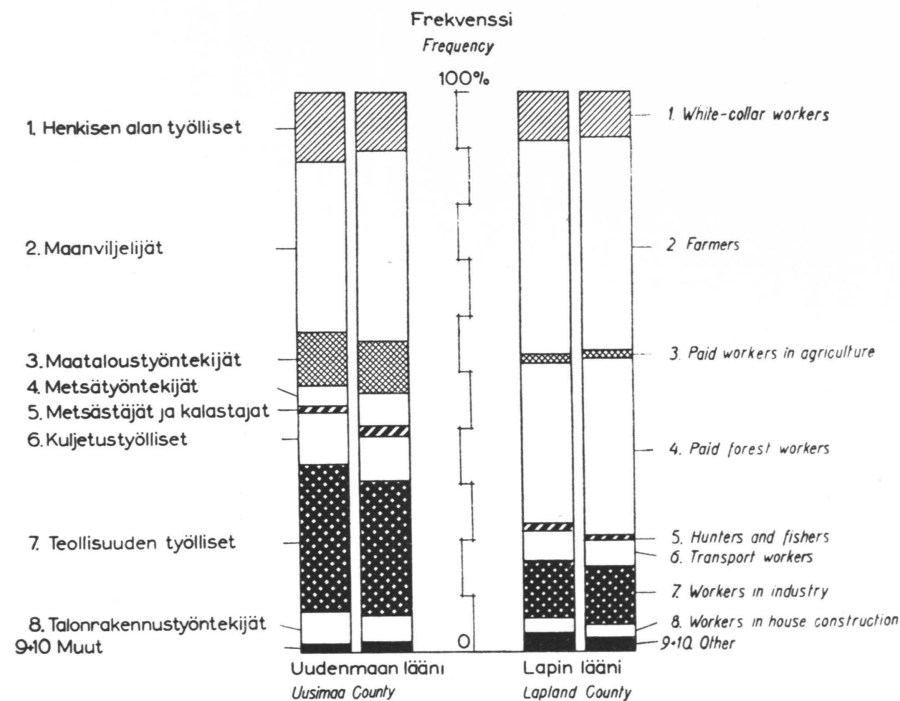
Ammattiryhmä Occupation group *	»Perus- joukko» »Popu- lation»	Näyte — Sample		
		Syntymä- päivä »Popu- lation» 1	Syntymä- päivä »Popu- lation» 15	Yhteensä Total
1. Henkisen alan työlliset	8.8	8.3	7.9	8.1
2. Maanviljelijät ja heidän maatilan töihin osallistuvat perheenjäsenensä. Turkkiläinten hoitajat, poronhoitajat ja puutarha-alan työlliset	38.2	37.3	39.5	38.2
3. Maataloustyöntekijät	1.5	1.4	1.5	1.5
4. Metsätalouden ja uiton työntekijät	28.6	30.8	32.2	31.5
5. Metsästäjät ja kalastajat	1.2	1.2	0.9	1.0
6. Kuljetustyölliset	5.5	5.2	4.1	4.7
7. Teollisuuden, käsityön, kaivoksien ja kivilouhimoiden työlliset	10.2	11.2	9.5	10.4
8. Talonrakennustyöntekijät	2.8	2.3	2.2	2.3
9. Katu-, rautatie-, tie-, vesi- ym. rakennusalan työntekijät	2.1	2.3	2.2	2.3
10. Ilmoittamattomat ammatit ja ammatit, joita ei voitu määrittellä	1.1	—	—	—
Yhteensä — Total %	100.0	100.0	100.0	100.0
Koehenkilöitä — Males in sample	.	1 040	935	1 975

* See foot-note, Table 13/p. 81.

3312. Syntymäpäivämerkintöjen kasautumisen aiheuttamat harhat

Koska käytetty poimintamenetelmä, perusjoukon 1. ja 15. päivänä syntyneiden henkilöiden valitseminen koehenkilöiksi vastaa todennäköisyysvalintaa, ei menetelmästä itsestään voi johtua harhoja. Sen sijaan poiminnan käytännöllinen suoritus, mainittuina kuukauden päivinä syntyneiksi merkittyjen tai omien tietojensa mukaan syntyneiden henkilöiden valinta näytteeseen saattoi, kuten edellä luvussa 2433 todettiin, johtaa siihen, että näytteestä tuli harhainen.

Sen seikan, että henkikirjoihin merkityt syntymäpäivät selvästi kasautuvat ainakin 1. päivän kohdalle erilaisten virhemerkintöjen takia, ei vält-



Kuva 4. Miesten ammattijakautuma sadanneksina. »Perusjoukon» ja näytteen vertailu. Uudenmaan ja Lapin läänit. Päättutkimus. Lähde: taulukot 13/s. 81 ja 14/s. 83.

Figure 4. Percentage distribution of males into occupation groups. Comparison between »population» and sample. Uusimaa and Lapland Counties. Main Survey. Source: Tables 13/p. 81 and 14/p. 83.

tämättä ole tarvinnut aiheuttaa saadussa näytteessä virheellisyyksiä, mikäli itse aineiston keräyksessä sanotut virheellisyydet korjattiin. Näin tuskin kuitenkaan aina tapahtui.

Niissä kokemusselostuksissa, jotka väestönlaskentatoimikuntien piti laskennan suorituksen jälkeen lähettää Tilastollisen päätoimiston väestönlaskentaosastolle, tuli olla mm. ilmoitus siitä tavasta, jolla T-miesten etsintä suoritettiin. Ohjeissa oli ko. etsintä kehoitettu suorittamaan niin, että vuoden 1950 henkikirjajäljennöksen mukaan luetteloiitiin siinä olevat T-miehet ja laskija vielä lomakkeita (varsinkin C-lomakkeita) kerätessään tarkkasi syntymäaikoja. Toimikunnista 404 antoi vastauksen kysymykseen T-miesten etsintätavasta. Näistä 147 ilmoitti sen tehdyn henkikirjan, 140 C-lomakkeiden, 87 henkikirjan ja C-lomakkeiden, 11 henkikirjan ja

kirkonkirjan avulla sekä 19 joko kaikkia edellä mainittuja apuvälineitä käyttäen, »syntymäajan perusteella» tai muilla tarkemmin erittelemättömillä keinoilla.

Kokemusselostuksissaan muutamat toimikunnat ovat nimenomaan maininneet työvoimatiedustelun tuottaneen hankaluutta sen tähden, että syntymäajat oli merkitty väärin henkikirjoihin ja niitä jouduttiin tarkistamaan kirkonkirjoista. Näyttää todennäköiseltä, että syntymäpäivämerkintöjen kasautumisilmiö ei ole ehkä koko kansaneläkelaitoksen aineistosta todetulla painollaan päässyt vaikuttamaan saatuun näytteeseen.

Koska on todettu, että virheellisten syntymäaikatietojen antaminen on yleisempää iäkkäissä ja vähän koulutietoja saaneissa väestöryhmissä, on todennäköistä, että kasautumisilmiö pyrki lisäämään näytteen vanhimpia ja vähemmän koulusivistystä saaneita ryhmiä sekä vastaavasti vähentämään nuoria ja koulusivistystä saaneita ikäluokkia. Edellisen luvun vertailevista tutkimuksista kävi ilmi, että näytteen ikäjakautumassa näyttää olevan lievä sanotun suuntainen poimintaharha. Myös näytteen ammattijakautumassa on, ainakin Uudenmaan läänissä, vertailun perusteella todettavissa, että maatalous- ja metsätyöntekijöitä on jonkin verran liian paljon. Tätä tulosta on kuitenkin pidettävä epävarmana. Suoritettujen tutkimusten perusteella ei voida varmasti päätellä, mitkä sanottujen mahdollisten virheiden syyt ovat. Voi olla, että kato ja poiminnasta johtunut kasautuminen ovat pyrkineet vaikuttamaan osittain samaan suuntaan.

332. Mittausarhat

Ainakin pinnallisesti asiaa tarkasteltaessa näyttää vuoden 1950 maaseudun työvoiman tutkimuksen kaikkein huomattavin virhelähde piilevän mittausmenetelmässä. Näin ollen ei ole ihmeteltävää, että kokemusselostuksissaan viisi väestönlaskentatoimikuntaa piti haastattelujen tuloksia epäluotettavina. Niinpä Viialan toimikunta leimasi ne »mielikuviuksen tuotteiksi», Kokemäen toimikunta taas väitti työvoimatiedustelun, kai oman kuntansa osalta, antaneen harhauttavan kuvan ja esitti, että »tutkimuksen edustavuuden parantamiseksi» olisi pitänyt valita erityyppisiä tiloja ja pyytää näillä työskenteleviä miespuolisia henkilöitä tekemään vuoden alusta lähtien muistiinpanoja suorittamistaan töistä. Käytetyn mittausmenetelmän luotettavuuteen on epäilemättä syytä suhtautua kriittisesti. Todettakoon kuitenkin, että asiaan vain pinnallisesti perehtynyt

Tau lukko 15. Miestyövoiman arkiajan käyttö sadanneksina Ranuan kunnassa v. 1950. Päättökimpuksen haastattelujen ja II esitutkimuksen kirjanpidon vertailu. H = haastattelututkimus, 73 koehenkilöä. K = kirjanpitolutkimus, 133 koehenkilöä, t = Studentin jakautuman t-arvo H-K:lle.

Table 15. Use of manpower in per cent, Ranua commune in 1950. Comparison between interviews in Main Survey and bookkeeping in Pilot Survey II. H = interview survey, 73 males in sample. K = bookkeeping survey, 133 males in sample. t = t-value for H-K in Student's distribution.

Ajankäyttoryhmä Activity group * (see p. 87)	Kuukausit — Month												Koko vuosi 1950 Total year				
	I		III		V		VII		IX		XI		H	K	H-K	t	
	H	K	H	K	H	K	H	K	H	K	H	K					
	Sadanneksia (\bar{p}_j) kaikista ¹ arkipäivistä — Per cent (\bar{p}_j) of all ¹ week-days																
1. Markkinapuun hakkuu ja ajo, metsänhoito, ansiotyö	62.4	63.2	29.9	29.2	4.9	6.2	2.6	4.0	6.6	6.5	21.6	22.0	20.2	20.9	—	0.7	0.4
2. Uitto, uittoväylien ja teiden rakennus	6.6	4.1	7.1	3.0	27.8	19.7	6.6	5.2	8.8	12.5	3.2	2.0	9.0	6.8	+	2.2	1.1
3. Maatalous, ansiotyö	1.5	0.6	0.8	1.0	2.2	5.0	8.6	7.2	5.8	5.5	1.1	2.6	3.6	4.0	—	0.4	0.4
4. Muu ansiotyö ja oma henkinen työ	7.9	8.7	11.7	10.3	13.9	14.4	17.9	13.6	11.9	18.6	13.5	14.2	13.1	13.9	—	0.8	0.2
5. <i>Ansioyöpanos ja oman henkisen työn panos yhteensä</i> (ryh- mät 1—4)	78.4	76.6	49.5	43.5	48.8	45.3	35.7	30.0	33.1	43.1	39.7	40.8	45.9	45.6	+	0.3	0.1
6. Metsästys, kalastus ja poronhoito	1.4	2.3	0.5	1.6	1.1	1.5	1.2	0.7	1.2	1.2	3.2	3.8	1.6	2.1	—	0.5	0.7
7. Markkinapuun hakkuu ja ajo, metsänhoito, oma työ	3.1	2.0	4.1	3.7	0.9	1.5	0.4	1.2	4.5	0.2	16.1	15.7	5.5	4.6	+	0.9	0.8
8. Talonrakennus, oma työ	—	0.1	3.4	2.0	3.5	5.0	3.8	4.2	10.6	5.6	6.4	0.8	5.3	3.4	+	1.9	2.1*
9. Uudisraivaus, oma työ	—	—	—	0.5	7.9	5.1	7.5	9.0	16.7	15.4	1.0	0.9	6.0	6.3	—	0.3	0.3
10. Muu maatalous, kotitarvepuun hakkuu ja ajo, oma työ	8.2	10.7	33.1	32.6	31.3	28.1	46.3	46.4	25.7	24.5	26.0	25.2	27.9	25.7	+	2.2	0.8
11. <i>Oman ruumiillisen työn panos yhteensä</i> (ryhmät 6—10)	12.7	15.1	41.1	40.4	44.7	41.2	59.2	61.5	58.7	46.9	52.7	46.4	46.3	42.1	+	4.2	1.1
12. <i>Työpanos yhteensä</i> (ryhmät 5—11)	91.1	91.7	90.6	83.9	93.5	86.5	94.9	91.5	91.8	90.0	92.4	87.2	92.2	87.7	+	4.5	2.3*
13. Vapaa-aika	2.7	0.7	1.1	1.4	1.1	1.2	1.1	0.9	3.2	2.2	1.7	1.5	1.7	1.8	—	0.1	0.3
14. Säiraus	1.3	0.8	1.3	3.3	1.5	2.3	2.0	2.4	1.8	2.0	2.9	1.8	1.9	2.4	—	0.5	0.8
15. Avoin työttömyys	0.3	1.7	3.4	7.5	1.0	7.3	0.5	2.5	0.9	3.3	0.7	5.5	1.4	4.5	—	3.1	3.9***
16. Muu hukka-aika	4.6	5.1	3.6	3.9	2.9	2.7	1.5	2.7	2.3	2.5	2.3	4.0	2.8	3.6	—	0.8	1.1
17. <i>Muu arkaika yhteensä</i> ¹ (ryh- mät 13—16)	8.9	8.3	9.4	16.1	6.5	13.5	5.1	8.5	8.2	10.0	7.6	12.8	7.8	12.3	—	4.5	3.5***
18. <i>Koko arkaika yhteensä</i> ¹ (ryhmät 12, 17)	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	+	0.8	1.0
19. Loppoaika	0.1	0.7	7.6	5.0	1.9	1.4	0.5	—	2.6	2.4	3.0	1.9	2.7	1.9	+	0.8	—

¹ Paitsi opiskeluun ja asevelvollisuuteen käytetyistä päivistä — Except days devoted to studies or military service.

ei tavallisesti osaa erottaa toisistaan tulosten luotettavuuden kannalta suhteellisen vaaratonta mittaustulosten satunnaispoikkeamaa mittausharhasta.

Vuoden ajankäytön tiedustelussa mahdollisesti syntyviä harhoja pyrittiin tutkimaan edellä selostetuilla esitutkimuksilla (luku 223). Tärkein mittausharhan tutkimus voitiin tehdä vasta päätutkimuksen tultua suoritetuksi Ranualla, vertaamalla siellä vuoden 1950 arkiajan käytöstä tehdyn kirjanpitolutkimuksen (II esitutkimus) tuloksia päätutkimuksen vastaaviin tuloksiin. Ranuan päätutkimus merkitään seuraavassa H:lla ja kirjanpitolutkimus K:lla.

K-tutkimuksen näyte (133 koehenkilöä) oli otos Ranuan henkikirjaan 1. 1. 1949 vv. 1885—1933 syntyneiksi merkityistä työkykyisistä miehistä. H-tutkimuksen seuraavassa käsiteltävä näyte (73 koehenkilöä) käsitti ne väestönlaskennassa haastatellut samanikäiset työkykyiset T-miehet, jotka olivat ilmoittaneet asuneensa Ranualla 31. 12. 1950. K-tutkimuksessa vastausten katon osuus oli 0 %, mutta H-tutkimuksessa oli $J_{max} = 24 \%$, samoin oli myöskin $J_{min} = 24 \%$.

Se seikka, että K- ja H-näytteet poimittiin osittain eri perusjoukoista, voi vähentää tulosten vertailukelpoisuutta varsin vähän. Virallisen tilaston (Väestötilastoa 1951, s. 13, 1952, s. 13 sekä Suomen tilastollinen vuosikirja 1950, s. 24 ja 1953, s. 21) perusteella laskettiin, että kunnasta muutti v. 1949 vain 2.1 % ja v. 1950 1.6 % kyseisten vuosien alussa todetusta miespuolisesta väestöstä. Vastaavat kuntaanmuuton sadannekset olivat v. 1949 1.7 % ja v. 1950 1.5 %. Sen sijaan verrattain suuri katon osuus heikentää vertailutulosten varmuutta, vaikka perusjoukko on suhteellisen tasalaatuinen. Kuitenkin on olemassa hyvin pieni todennäköisyys, että perusjoukkojen eroavuus ja katon aiheuttama H-tutkimuksen harhaisuus voisivat yhdenmukaistaa H- ja K-tutkimusten tuloksia. Mainittuja seikkoja ei sen tähden seuraavassa oteta huomioon.

Myös kirjanpitolutkimuksessa voi syntyä mittausharhoja (vrt. lukua 223). Koska varmempaan aineistoon ei ole käytettävissä, pidetään seu-

* 1. Felling and hauling of wood for sale, silviculture, paid work. — 2. Floating, construction of roads and water-ways. — 3. Agriculture, paid work. — 4. Other paid work and unpaid white-collar work. — 5. Total input of paid labour and unpaid white-collar labour (groups 1—4). — 6. Hunting, fishing and reindeer raising. — 7. Felling and hauling of wood for sale, silviculture, unpaid work. — 8. House construction, unpaid work. — 9. Land clearing, unpaid work. — 10. Other agricultural work, felling and hauling of wood for domestic use, unpaid work. — 11. Total input of unpaid manual labour (groups 6—10). — 12. Total labour input. — 13. Holidays. — 14. Illness. — 15. Unemployment. — 16. Other wasted days. — 17. Total non-working week-days (groups 13—16). — 18. Total activity (groups 12, 17). — 19. Days underemployed.

raavassa K-tutkimuksen mittaustuloksia harhattomina. Käsittelyltään molempia tutkimuksia pidetään harhattomina.

Mainittujen seikkojen lisäksi H- ja K-tutkimusten tulokset voivat poiketa toisistaan paitsi H-tutkimuksen mittausharjojen takia myös molempien tutkimusten mittauksen, poiminnan ja käsittelyn satunnaispoikkeamista johtuen. Kaikkien niiden vaikutus on samanaikaisesti mitattavissa estimaatin keskihajonnalla (HANSEN ym. 1953, s. 91).

Taulukkoon 15/s. 86 on H- ja K-näytteistä laskettu estimaatit \bar{p}_j erälle vertailukelpoisille ajankäyttöryhmille (j) sekä kuukausittain (vain joka toinen kuukausi) että koko vuodelle 1950. Taulukon perusteella on piirretty kuva 5. Jotta poiminnan satunnaisvaihtelun aiheuttama ero tuloksissa saataisiin mahdollisimman vähäiseksi, laskettiin luvut \bar{p}_j seuraavalla tavalla.

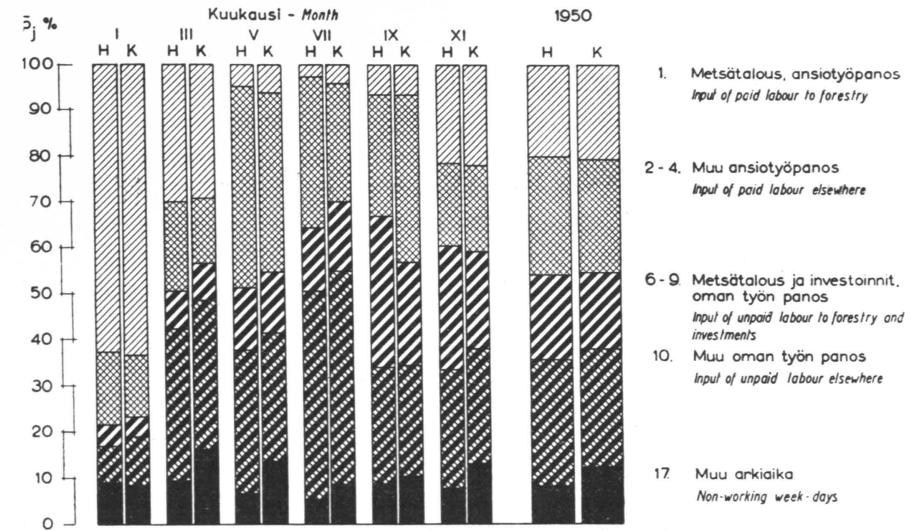
K-näytteessä on 8 koehenkilöä, jotka olivat koko tutkimusvuoden tai ainakin pääosan siitä opiskelemassa tai asevelvollisina. H-näytteeseen tällaisia koehenkilöitä ei tullut yhtään. Koska tämän seikan aiheuttama ero H- ja K-tutkimusten \bar{p}_j -estimaateissa olisi johtunut poimintavaihtelusta (tai vastausten kadosta) eikä ilmeisestikään mittausharhasta, jätettiin taulukon \bar{p}_j -lukuja laskettaessa huomioon ottamatta ajankäyttöryhmät asevelvollisuus ja opiskelu (K-näytteestä ao. koehenkilöitä ei siis poistettu). Suhdeluvut \bar{p}_j laskettiin siten kaikkien muiden arkipäivien summasta.

H- ja K-tutkimusten näin laskettujen koko vuoden \bar{p}_j -lukujen erotuksille laskettiin kaavalla 1/s. 37 sekä kaavalla

$$t^2 = \frac{(\bar{p}_{Hj} - \bar{p}_{Kj})^2}{\frac{s_{\bar{p}_{Hj}}^2}{2} + \frac{s_{\bar{p}_{Kj}}^2}{2}}$$

STUDENTIN jakautuman t -arvot, jotka merkitsevyyystunnuksineen on merkitty taulukkoon 15/s. 86 (vrt. esim. ALAMERI ja PÖYHÖNEN 1953, s. 61). Laskelmissa käytettiin 50 koehenkilön arvottua alanäytettä kummastakin tutkimuksesta.

Testin tulos osoittaa, että ainoastaan ryhmässä 17 (muu arki-aika) sekä sen alaryhmässä 15 (työttömyysaika) on erittäin merkitsevä ero ($P < 0.001$). Ryhmissä 12 (työpanos yhteensä) sekä 8 (omat talonrakennustyöt) on melkein merkitsevä ero ($P < 0.05$). Kaikissa muissa taulukkoon 15/s. 86 merkityissä ryhmissä ero ei ole merkitsevä ($P > 0.05$). Mainittakoon, että jos vastaavat \bar{p}_j -estimaatit lasketaan kaikkien arkipäivien summasta, tulos



Kuva 5. Miestyövoiman arki-aajan käyttö sadanneksina Ranuan kunnassa v. 1950. Pää-tutkimuksen haastattelujen (H) ja II esitutkimuksen kirjanpidon (K) vertailu. Lähde: taulukko 15/s. 86. Ryhmien numerot viittaavat taulukkoon.

Figure 5. Use of manpower in per cent, Ranua commune in 1950. Comparison between interviews in Main Survey (H) and bookkeeping in Pilot Survey II (K). Source: Table 15/p. 86. Numbers of activity groups refer to the foot-note in the table.

on muuten sama, mutta ryhmässä 12 erotus tulee erittäin merkitseväksi ja taulukosta puuttuvaan ryhmään »opiskelu ja asevelvollisuus» tulee merkitsevä ero.

H- ja K-tutkimuksista taulukkoon 15/s. 86 laskettujen \bar{p}_j -arvioiden erotukset pääsevät selvemmin näkyviin, jos ne lasketaan sadanneksina K-näytteen ao. sadanneksista. Taulukkoon 16/s. 90 on laskettu mainitut erotussadannekset taulukon 15/s. 86 pääryhmissä 5 (»ansiotyöpanos»), 11 (»oman työn panos»), 12 (»koko työpanos») sekä 17 ja 19 yhdistettyinä (»muu arki-aika»). Huomattakoon, että nimitykset ovat epätarkkoja ryhmän 12 nimeä lukuun ottamatta, ne ilmaisevat kuitenkin ryhmän pääasiallisen sisällön. Taulukkoon 16/s. 90 on otettu myös edellisestä taulukosta puuttuvat kuukaudet.

Koko vuoden tuloksissa on koko työpanos saatu H-tutkimuksessa 5 % suuremmaksi kuin K-tutkimuksessa. Tämä ero keskittyy »oman työn panoksen» ryhmään, jossa se on + 10 %. Toisaalta on »muun arki-aajan»

Taulukko 16. Ranuan haastattelu- ja kirjanpitolukemusten tulosten vertailu. Haastattelu- ja kirjanpitolukemuksesta saatuun sadannesteen (\bar{p}_j) erotus sadanneksina jälkimmäisistä. Päättökäytös ja esitutkimus II.

Table 16. Comparison between results of interview survey and bookkeeping survey at Ranua. Differences between percentages (\bar{p}_j) of interview survey and bookkeeping survey, expressed in per cent of latter. Main Survey and Pilot Survey II.

Ajankäyttöryhmä * Activity group *	Kuukausi — Month												Vuosi Year 1950
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	
5. »Ansiotyöpanos» »Input of paid labour»	+ 2	+15	+14	-13	+ 8	+16	+19	-10	-23	-11	- 3	- 6	+ 1
11. »Oman työn panos» — »Input of unpaid labour»	-16	-16	+ 2	+30	+ 8	+ 2	- 4	+17	+25	+19	+14	+14	+10
12. »Koko työpanos» Total labour input	- 1	+ 7	+ 8	+ 9	+ 8	+ 9	+ 4	+ 6	+ 2	+ 5	+ 6	- 1	+ 5
17+19. »Muu arki-aika» »Non-working days»	± 0	-47	-19	-35	-44	-55	-34	-44	-13	- 7	-18	+34	-26

* Ryhmien numerot viittaavat taulukkoon 15/s. 86. — Group numbers refer to Table 15/p. 86.

määräksi saatu H-näytteestä 26 % pienempi luku kuin K-näytteestä. Kuukausittaisissa tuloksissa on pääosiltaan samansuuntaiset erot, vaikka niiden suuruus vaihtelee. Näyttää siltä, että virhe »ansiotyöpanoksessa» on +-merkkinen ansiotöiden suhteellisen määrän ollessa suuri ja päinvastoin — -merkkinen niiden määrän ollessa pieni. Samaa osoittaa myös »oman työn panos». Vaikkakaan tämä ilmiö ei ole poikkeukseton, se voinee viitata siihen, että haastatteluvastauksissaan koehenkilöt helposti pyrkivät liioitellen ilmoittamaan suuret luvut liian suuriksi ja pienet luvut liian pieniksi. Sama ilmiö voinee olla myös selityksenä sairaus- ja työttömyysajoissa havaittavaan suureen eroon. Koska tämä aika kirjanpitolukemusten mukaan esiintyy koehenkilöiden työvuoden satunnaisina hajapäivinä, ne helposti unohtuvat. Arvattavasti koehenkilö pyrkii itsetiedottomasti kaunistelemaan ajankäyttöään eikä mielellään myönnä jälkepäin kovin suuren osan arkiajasta menneen hukkaan. Mutta tämä *pienien ryhmien kato* voinee esiintyä myös muissa pienissä ja samalla hajallisissa ajankäyttö-

ryhmissä, joista hyvänä esimerkkinä on metsästys ja kalastus, milloin niitä ei harjoiteta varsinaisena ammattina. Tutkimus osoittaa tämän ryhmän tulleen H-näytteessä miltei kaikkina kuukausina pienempi kuin K-näytteessä. Ranualla oli hyvin vähän ammattimetsästäjiä ja -kalastajia (sen sijaan lukuisia ammattimaisia poronhoitajia).

Ranualla suoritetuissa esitutkimuksissa tehtiin se virhe, että kirjanpitolukemusten koehenkilöistä ei kerätty vuoden 1951 alussa järjestelmällisesti haastattelumateriaalia vuoden 1950 ajankäytöstä. Tosin sellaisen tutkimuksen tulos, jos se olisi ollut hyvin yhtäpitävä kirjanpitolukemusten tulosten kanssa, olisi voitu tulkita osoitukseksi siitä, että koehenkilöt kirjanpitolukemusten ansiosta hyvin muistivat ajankäyttönsä rakenteen. On myös sängän luultavaa, että useat kirjanpidon pitäjät olisivat kieltäytyneet antamasta vastauksia tämännäytteisessä saman asian uudelleen kyselyssä. Toisaalta eroavuudet olisivat ehkä antaneet mahdollisuuden tarkemmin analysoida haastatteluvastauksien virheitä. Sattumalta oli H- ja K-näytteissä 5 identtistä koehenkilöä. Vertaamalla em. tavalla heidän antamiaan x_{ij} -lukuja H- ja K-näytteissä voitiin todeta, että erot olivat yhtä koehenkilöä lukuun ottamatta moninkertaiset taulukon 15/s. 86 erotuksiin verrattuna, mutta yleensä samansuuntaiset kuin siinä.

Vaikka haastattelussa saadut vastaukset ilmeisesti yksityistapauksissa voivat poiketa todellisista arvoista hyvinkin paljon, tämä *poikkeaminen näyttää edellä mainittua pienten ryhmien katoa lukuun ottamatta olevan laadultaan satunnaista*. Kun Ranuan koehenkilöt miltei poikkeuksetta voidaan lukea vaikeimmin haastateltavien luokkaan ja otetaan huomioon erotusten varmuusalue, voitaneen H- ja K-tutkimusten tuloksia pitää ainakin välttävästi yhdenmukaisina. Ihme kyllä, ei voitu havaita H-tutkimuksen tulosten havaintokauden alkukuukausina olleen harhaisempia kuin loppuvuoden kuukausinakaan. Suurissa näytteissä todetut eroavuudet ilmeisesti pyrkisivät yhäkin pienemään. *Pienten ryhmien kato on kuitenkin otettava huomioon ajankäytön haastattelututkimuksen tuloksia tulkittaessa*. Tämä koskee ennen kaikkea ryhmää työttömyys.

Myös kulutustutkimuksissa on Kansainvälisen työtoimiston ilmoituksen mukaan (Methods ... 1949, s. 13) verrattu kirjanpito- ja haastattelututkimuksia toisiinsa »muutamissa koetutkimuksissa». Tosin näistä tietävästi on julkaistu vain Norjan Förbrukundersökelse (1953). Siinä käytettiin vertailussa kahta mainituin tavoin mitattua saman perusjoukon rinnakkaista näytettä. Tuloksia pidettiin tyydyttävän yhdenmukaisina. Niiden eroavuutta ja sen syytä ei mainitussa tutkimuksessa analysoitu.

Jos muistiin perustuvaa, haastatteleamalla tapahtuvaa perheen koko

vuoden kulutuksen mittaamista verrataan samalla tavalla suoritettavaan yhden henkilön ajankäytön arviointiin, jossa päivä on mittayksikkönä, voidaan tuskin päätyä muuhun tulokseen, kuin että edellinen on ainakin yhtä suuressa määrin muistia rasittava kuin jälkimmäinen. Huomattakoon, että perhebudjettia mitattaessa myös kokonaiskulutuksen määrä on useimmiten muistin varassa, kun taas ajankäyttötutkimuksessa vuoden arkipäivien summa on määrätty ja kaikilla koehenkilöillä sama (pyhätyöpäivät tosin muodostavat vähäisen poikkeuksen).

Pääaineiston keräyksessä käytetyn mittausten menetelmän luonteesta johtuu, että haastattelijan työskentelytapa ratkaisevasti vaikuttaa mittaus-tulosten luotettavuuteen. Työvoimatiedustelun suorittaneet väestönlaskennan laskijat valittiin alun perin opettajista, opiskelijoista, oppikoululaisista ym. kohtalaisen opillisen sivistyksen saaneista henkilöistä, joiden joululoman aikaan väestönlaskenta hyvin sopi alkuperäisen suunnitelman mukaan. Lomakkeiden painatuksen viivästymisestä johtui, että monin paikoin näiden valittujen laskijoiden, joille oli ehditty antaa välttämättömin koulutuskin, oli palattava oppilaitoksiinsa ennen laskennan suorittamista. Heidän tilalleen jouduttiin ottamaan laskijoita, jotka eivät kokemuselostusten mukaan aina kyenneet selviytymään tehtävistään täysin tyydyttävästi. Kun T-lomakkeen täyttäjältä vaaditaan parhaan mahdollisen tuloksen saavuttamiseksi suuri pätevyys, tämä oli luonnollisesti omiaan heikentämään tulosten luotettavuutta. Samaan suuntaan vaikutti myös laskijoiden suuri lukumäärä (pyörein luvuin 10 000), joskin samalla täytettävien T-lomakkeiden luku laskijaa kohti siten tuli pieneksi (keskim. 5 lomaketta). Se esti haastatteluihin kyllästymisen, ja laskijoista johtuvat satunnaiset mittausvirheet pääsivät hyvin tasoittumaan. Toisaalta väestönlaskijoiden hoidettavina olleiden erilaisten tutkimusten suuri lukumäärä oli omiaan aiheuttamaan laskijoiden kyllästymistä tehtäviinsä. Seitsemän toimikuntaa on ilmoittanut myös T-miesten keskuudessa todetun väsymystä moniin tiedusteluihin ja siitä johtuvaa vastahakoisuutta. Viisi toimikuntaa on ilmoittanut T-miesten suhtautuneen epäluuloisesti työvoimatiedusteluun ja niin muodoin olleen haluttomia antamaan pyydettyjä tietoja.

Laskentatoimikuntien ilmoituksista voi päätellä T-miesten joissakin tapauksissa itse täyttäneen lomakkeensa. On tosin vaikeaa arvioida, onko tämä seikka vaikuttanut tuloksiin heikentävästi vai päinvastoin parantavasti. Jokin toimikunta on nimenomaan moittinut sitä ohjetta, jonka mukaan T-lomakkeita ei saanut etukäteen jakaa T-miehille.

Mikäli täytettyjen T-lomakkeiden perusteella pystyy arvostelemaan niiden täyttäjien pätevyyttä, näyttää valtaosa laskijoista suhtautuneen

tehtävänsä vakavasti ja huolella. Varsinkin maan itä- ja pohjoisosien kunnissa on kuitenkin ollut huomattava joukko kirjallisiin töihin varsin tottumattomia laskijoita, joiden jälkiä lomakkeiden tarkistusmenettelyssä jouduttiin korjaamaan. Tulosten kannalta on valitettavaa, että laskijat ovat olleet keskimäärää heikompia juuri niillä alueilla, joissa haastattelun suorittaminen oli yleensä vaikeaa. Toisaalta ei voida pitää selvänä, että koulua käymättömät laskijat olisivat olleet vähemmän tunnollisia tehtävissään kuin muut, vaikka lomakkeisiin onkin tullut muodollisia virheitä ja puutteellisuuksia.

Myös T-lomakkeen ajankäyttöryhmissä ja niiden nimikkeissä oli sellaisia tulkintavaikeuksia ja virheitä, jotka olivat omiaan lisäämään mittausvirheitä. Niinpä raakapuun autokuljetuksen ja lastauksen erottaminen omaksi ryhmäkseen tuskin oli onnistunut toimenpide, koska samat henkilöt varsin usein harjoittavat sekä raakapuun että muun tavaran autokuljetusta samoina päivinäkin. Näin ollen mainittua ryhmää tuloksissa on pidettävä muita huomattavasti epävarmempana. Sama koskee julkisten rakennustöiden erottamista asuin- ja talousrakennustöistä, mikä näyttää tuottaneen laskijoille ja haastateltaville liian suuria vaikeuksia. Vielä näyttää (tuloksista päätellen) omassa metsässä suoritettuja metsänhoidollisia töitä olleen vaikea pitää erillään kotitarvepuun hakkuusta ja ajosta. Yleensä näyttää »kaatoryhmän» tapainen nimike, kuten »muut työt omassa metsässä», »muu aika, jota ei ole käytetty työhön» jne., houkuttelleen liian paljon laskijoita. Viimeksi mainituissa ryhmässä vaadittu kyseisen ajankäytön selitys toisaalta auttoi aineiston käsittelyssä siirtämään päivät oikealle rivilleen. Edellä mainituista ryhmien epävarmoista rajoista huolimatta ei ole katsottu olevan syytä yhdistellä niitä tulostaulukoissa, koska ne antanevat käsityksen näiden ryhmien suuruusluokasta.

333. Käsittelyerheet

Luvussa 322 selostettu aineiston käsittely sekä siihen liittyvä näytteestä laskettujen tunnuslukujen suurentaminen (luku 323) ei luonnollisesti voinut tapahtua virheitä, vaikka niiden välttämiseksi ryhdyttiin lukuisiin varokeinoin. Päinvastoin on todettava, että kaikissa käsittelyvaiheissa huolellisimpienkin työntekijöiden ja uusimpienkin tilastokoneiden havaittiin tekevän virheitä. Käsittelymenetelmien monimutkaisuus oli omiaan vielä lisäämään virhemahdollisuuksia. Monilaatuisten käsittelyvirheiden syihin ja niiden tarkistus- ja korjausmenetelmiin ei tässä ole

aihetta yksityiskohtaisesti syventyä. Todettakoon vain, että yleensä ei havaittu virheiden olevan laadultaan systemaattisia.

Poikkeuksen muodostavat ne käsittelyvaiheet, joissa T-lomakkeen päivämerkinnät merkittiin käsin reikäkortteille, laskettiin yhteen koneellisesti, kerrottiin suurennustekijällä ja jälleen laskettiin yhteen taulukoiksi. Näissä käsittelyissä päivien määrä pyrki ennemmin tulemaan liian pieni kuin liian suuri. Suoritetut tarkistuslaskelmat ovat osoittaneet, että syntyneet virheet peittyivät pyörityksissä ja tasoitusmenettelyssä, jossa pyöristetyt yhteenlaskettavat luvut tasoitettiin niin, että niiden summa tuli samaksi kuin pyöristämättömistä luvuista laskettu loppusumma. Tasoitukset suoritettiin noudattamalla periaatetta, että tasoitusvirheen suhteellinen suuruus tulee silmämääräisesti arvioiden mahdollisimman pieni.

334. Satunnaisvaihtelu

3341. Työvoiman määrän varianssi

Sekä poiminnan, mittauksen että käsittelyn satunnaisvaihtelun vaikutus otoksesta laskettujen arviolukujen varmuuteen on mahdollista todeta matemaattisin keinoin laskemalla estimaattien keskihajonnat (keskivirheet). Edellytyksenä on, että kunkin näyteyksikön tunnuslukujen satunnaispoikkeamat ovat riippumattomia toisten näyteyksiköiden satunnaispoikkeamista (HANSEN ym. 1953, s. 91). Päättökäytöksessä aineiston kerääjien lukumäärä oli lähimain 10 000, ja T-lomakkeisiin tehtävät korjaukset tarkisti valtaosalta yksi henkilö. Näin voidaan mainittua vaatimusta pitää käytännöllisesti katsoen täytettynä. Kun päätökäytöksen näytettä on pidettävä otosta vastaavana, näytteestä lasketuille tunnusluvuille johdetut poimintavarianssin kaavat ilmaisevat näin ollen samalla kaiken satunnaisvaihtelun vaikutuksen tunnuslukujen varmuuteen. Seuraavassa käytetään arviolukujen varianssin laskemiseksi *osittamattoman*, yksiasteisen otoksen menetelmiä. Kun näytteen suurentaminen suoritettiin lääneittäin, metsänhoitolautakuntien toiminta-alueittain ja työvoimapiireittäin perusjoukon ja näytteen henkilömäärien suhteella, on näytettä pidettävä ositettuna otoksena. Täten saadaan osittamattoman otoksen kaavoilla koko valtakunnan arvioluvuille jonkin verran liian suuret varianssit (ja keskivirheet).

Maaseudun työvoiman tutkimuksen näytteestä laskettiin kahta perustyyppiä olevia tunnuslukuja. Laskentayksikkönä oli joko perusjoukkoon

kuuluva *henkilö* (näytteessä koehenkilö) tai *miespäivä* (em. henkilön arkipäivä v. 1950).

Ensiksi mainituissa tutkimustuloksissa, joista taulukko 17/s. 97 on esimerkkinä, voidaan estimaattien varmuutta arvioitaessa käyttää kvantitatiivisen analyysin menetelmiä. Tässä analyysissä käytetään seuraavia merkkejä:

n koko näytteen koehenkilöiden lukumäärä ¹ T_h jollakin kunnan tai kuntaryhmän muodostamalla alueella.

N vastaavan perusjoukon miesten lukumäärä.

$K = \frac{N}{n}$. Lopullinen suurennustekijä kyseiselle alueelle.

m määrättyllä yhteisellä ominaisuudella varustettujen koehenkilöiden lukumäärä.

μ m -ryhmän alaryhmä näytteessä siten, että μ koehenkilöllä on myös jokin toinen yhteinen ominaisuus.

$m' = Km$. Näytteestä laskettu arvioluku.²

$\mu' = K\mu$. Näytteestä laskettu arvioluku.

$$\varphi = \frac{m}{n} = \frac{m'}{N}$$

$$\psi = \frac{\mu}{m} = \frac{\mu'}{m'}$$

Estimaattien m' ja φ suhteellinen varianssi saadaan mitättömän harhaisena kaavasta (esim. HANSEN ym. 1953, s. 125)

$$v_{m'}^2 = (1 - \varphi) / n\varphi = (1 - \varphi) / m = v_{\varphi}^2 \quad (5)$$

Tuloksissa esitetään kuitenkin yleisimmin estimaatteja μ' ja ψ . Ensiksi mainitun suhteellinen varianssi saadaan (mt. s. 126) hyvänä likiarvona kaavasta

¹ Oikeampaa olisi asettaa $n = T_h + T_a$. Virhe on merkityksetön.

² Todellisuudessa suurentaminen kyseisissä tuloksissa suoritettiin lääneittäin, joten oikeammin olisi merkittävä $m' = \sum_r^g K_r m_r$, jossa K_r on r :nnen läänin suurennustekijä, m_r määrättyllä ominaisuudella varustettujen näyteyksiköiden lukumäärä samassa läänissä ja g läänien lukumäärä näytteessä. Yksinkertaisuuden vuoksi seuraavat kaavat johdetaan niin, että K :lla tarkoitetaan kulloinkin kysymyksessä olevan alueen (kuntien näytteillä painotettua) kuntien suurennustekijöiden keskiarvoa. Näin syntyvä virhe on merkityksetön.

$$I \quad v_{\mu'}^2 = v_m^2 + (1 - \psi) / m \psi, \quad (6)$$

joka on taulukon 17/s. 97 tyyppisissä tuloksissa hyvin käyttökelpoinen. Absoluuttiset varianssit ovat laskettavissa kaavoista (on huomattava, että $v_{\mu'}^2$, ei ole $= v_{\psi}^2$)

$$I \quad s_{\mu'}^2 = v_{\mu'}^2 \mu'^2 \quad (7)$$

$$I \quad s_{\psi}^2 = \psi (1 - \psi) / m \quad (8)$$

Viimeksi esitetty kaava antaa HANSENIN ym. mukaan riittävän harhattoman tuloksen, jos $\mu \geq 20$ (mt. s. 125). Varovaisuuden vuoksi päättökäytön tuloksissa kuitenkin asetettiin mainitun kaavan käytölle vaatimukset $m \geq 100$ ja $\mu \geq 35$ (vrt. mt. s. 131 ja COCHRAN 1953, s. 41).

Edellä esitetyt kaavat perustuvat olettamukseen, että muuttujien φ ja ψ todennäköisyysjakautuma on binomiaalinen. Olettaen, että se on POISSONIN jakautuma, saadaan (WEATHERBURN 1947, ss. 47—48)

$$II \quad v_{\mu'}^2 = K / \mu' \quad (6 a)$$

$$II \quad s_{\mu'}^2 = K \mu' \quad (7 a)$$

$$II \quad s_{\psi}^2 = \psi / m \quad (8 a)$$

Olettamus ψ -muuttujan binomiaalisesta todennäköisyysjakautumasta on oikeampi, sillä kun $\psi = 1$, on luonnollisesti $s_{\psi}^2 = 0$. Kuitenkin antavat ryhmän II hyvin helppokäyttöiset kaavat likimäärin saman tuloksen kuin ryhmän I kaavat, jos kaavaa 7 a käytetään esim. vain, kun $\psi \varphi \leq 0.2$ ja kaavaa 8 a, jos $\psi \leq 0.2$. Näillä edellytyksillä tulee II ryhmän kaavoilla korkeintaan 11.8 % suurempi keskivirhe kuin ryhmän I kaavoilla. Tämä näkyy siitä, että

$$II \quad s_{\mu'} / I \quad s_{\mu'} = \sqrt{1 / (1 - \psi \varphi)} \quad \text{ja}$$

$$II \quad s_{\psi} / I \quad s_{\psi} = \sqrt{1 / (1 - \psi)}$$

Taulukko 17. Metsä- ja uittotyöntekijöiden jakautuminen metsä- ja uittotöiden ansiotyöpanoksen suuruusluokkiin v. 1950. Koko maaseutu. Päättökäytön tutkimus.

Table 17. Distribution of paid forest and floating workers according to input of paid labour to forestry and floating in 1950. All rural districts. Main Survey.

Työpanos, päiviä Labour input in days	Näytteessä miehiä Males in sample	Perusjou- kossa miehiä ¹ Males in popula- tion ¹	%		Keskivirhe ² Standard error ² $S_{\mu'}$		
			φ	ψ	I	II	III
1—50	8 640	149 317	18.4	48.2	1 454	1 602	1 154
51—100	5 084	87 479	10.8	28.3	1 164	1 228	1 039
101—150	2 355	40 365	5.0	13.1	816	834	777
151—200	1 078	18 388	2.3	6.0	558	564	546
201—250	512	8 756	1.1	2.9	388	390	383
250+	266	4 564	0.6	1.5	281	282	278
Yhteensä — Total							
Näytteen alaryhmä — Subgroup of sample. m	17 935	.	} 38.0	100.0	.	.	.
Perusjoukon alaryhmä Subgroup of population m'	.	308 869			1 822	2 307	0
Koko näyte — Total sample n	46 853	.	} 100.0	.	.	.	
Koko perusjoukko — Total population .. N	.	812 726		0	3 743	.	

Ensiksi mainittu yhtälö on helpoimmin todistettavissa sijoittamalla sen vasemmalle puolelle I $s_{\mu'} = \sqrt{\mu'^2 (N - \mu') / n \mu'}$, joka saadaan kaavasta 5/s. 95, koska μ' on m' -käsitteen erikoistapaus. Kuten näkyy, antavat ryhmän II kaavat aina suuremman keskivirheen kuin ryhmän I kaavat.

Silloin, kun mainitut vaatimukset eivät ole täytetyt, mutta v_m^2 on niin

¹ Suurennus on suoritettu lääneittäin — Raising computed for each county separately.

² I Kaavasta 7/s. 96 — According to formula 7/p. 96.

II Kaavasta 7a/s. 96 — According to formula 7a/p. 96.

III Kaavasta 7b/s. 96 — According to formula 7b/p. 96.

pieni, että se voidaan jättää huomioon ottamatta ($m \sim n$), saa kaava 7/s. 96 likiarvomuodon

$$\text{III} \quad s_{\mu'}^2 = \mu'^2 (1 - \psi) / m \psi, \quad (7 \text{ b})$$

joka antaa aina liian pienen keskivirheen kun $m < n$.

Taulukkoon 17/s. 97 on esimerkin vuoksi laskettu μ' -estimaattien keskivirheet kaikilla näillä kolmella menetelmällä (kaavoista 7, 7 a/s. 96 ja 7 b). Tulos osoittaa, että kaikki kaavat antavat tässä tapauksessa verrattain yhtäpitävän tuloksen. Koska keskivirheen yliarviointi on suositeltavampaa kuin sen aliarviointi, on yksinkertaisin kaava 7 a tässä ja vastaavanlaisissa hyvin usein esiintyvissä tapauksissa erittäin käyttökelpoinen.

3342. Työpanoksen ja muun arkiajan käytön määrän varianssi

Vuoden 1950 maaseudun työvoiman tutkimuksen tulosten toisen pääryhmän muodostavat ne, joissa esitetään perusjoukon määrättyllä ominaisuudella varustetun osan m' arkipäivien summan z' jakautuminen ajankäyttöryhmiin (j) joko absoluuttisina lukuina x'_j tai suhdelukuina $x'_j/z' = \bar{p}_j$ (merkkijärjestelmä selostettiin luvussa 222). Esimerkkinä tämän tyyppisistä tuloksista on taulukko 18/s. 99. Tässäkin tapauksessa on kysymys kvalitatiivisesta analyysistä. Ryhmä miespäiviä z' jakautuu osaryhmiin x'_j , joilla kullakin määrätty ominaisuus j on yhteinen. Tässä ei sellaisinaan voida käyttää edellä esitettyjä kaavoja (sijoittamalla niihin merkin μ' paikalle x'_j , merkin m' paikalle z' jne.). Tämä johtuu siitä esim. YATESIN (1949, s. 195) mainitsemasta seikasta, että kvalitatiivisen analyysin varianssikaavat pätevät vain siinä tapauksessa, että ne yksiköt (tässä miespäivät), joiden ominaisuudesta on kysymys, ovat myös näytteen poimintayksiköitä. Koska näytteen poimintayksikkönä päätutkimuksessa oli mies eikä miespäivä, ei tätä vaatimusta ole täytetty. YATESIN mukaan on tunnuslukujen varianssit tällöin laskettava kvantitatiivisen analyysin menetelmin (esim. kaavasta 1/s. 37).

Viimeksi mainittu menetelmä on monessa suhteessa työläs käyttää testattaessa x'_j - ja \bar{p}_j -tunnuslukujen ja niiden johdannaisten keskinäisten erotusten merkitsevyyttä. Vaikka kaavaa 1/s. 37 käytettäessä tyydyttäisiin suhteellisen pieniin alaotoksiin, näiden otosten poimiminen koko näytteestä ja varianssien laskeminen on paljon työtä ja kustannuksia vaa-

Taulukko 18. Metsä- ja uittotyövoiman arkiajan käyttö v. 1950. Lapin metsänhoitolautakunnan toiminta-alue. Päättökäytös.

Table 18. Use of forest and floating manpower in 1950. District of Lapland Forestry Board. Main Survey.

Ajankäyttöryhmä — Activity group	x'_j 1000 mies- päivää man-days in thousands	\bar{p}_j %	c	Keskivirhe ¹ Standard error ¹ $s_{x'_j}$ 1000 miespäivää man-days in thousands	
				I	II
1. Maatalous — Agriculture	2 406	29.9	0.75	58	84
2. Metsästys, poronhoito ja kalastus — Hunting, reindeer raising and fishing	263	3.3	0.75	27	28
3. Metsätalous — Forestry	2 177	27.1	0.40	48	43
4. Raakapuun kaukokuljetus — Trans- portation of roundwood	338	4.2	0.55	23	23
5. Talonrakennus — House construction	645	8.0	0.75	43	44
6. Teiden, rautateiden ja uittoväylien rakennus — Construction of roads, railways and water ways	468	5.8	0.55	27	27
7. Käsityö ja pienteollisuus — Domes- tic handicraft and minor industries	96	1.2	0.75	17	17
8. Muu työpanos — Other labour input	390	4.9	0.75	34	34
9. Työpanos yhteensä — Total labour input	6 783	84.4	0.55	102	104
10. Muu arki aika — Non-working week- days	1 258	15.6	0.55	47	45
Arki aika yhteensä — Total activity (z')	8 041	100.0	0	113	0

tiva toimenpide, koska tuloksissa esitetään suuri määrä mainittuja tunnuslukuja. Myöskin näin laskettujen keskivirheiden esittäminen tulosten yhteydessä on vaikeaa sekä painatusteknisistä että kustannussyistä. Tästä johtuen on seuraavassa pyritty kehittämään menetelmä, jolla myös x'_j - ja \bar{p}_j -estimaateille (samoin kuin μ' -, φ - ja ψ -estimaateillekin) asianomaisista tunnusluvuista itsestään ilman perusaineiston apua ovat laskettavissa keskivirheet käyttökelpoisella varmuudella.

¹ I Kaavasta 13/s. 115 — According to formula 13/p. 115.

II Kaavasta 13a/s. 115 — According to formula 13a/p. 115.
 $n = 2117$, $m = 1524$, $K = 17.5$, $a = 301.5$.

Menetelmän johto perustuu seuraavaan toteamukseen. Oletetaan että on kysymyksessä jonkin alueen (kunnan tai kuntaryhmän) koko perusjoukolle lasketut tunnusluvut x'_j ja \bar{p}_j , että z_i on kaikilla koehenkilöillä sama a , ja lisäksi että x_{ij} on kaikilla koehenkilöillä $= a$ (jokaisessa T-lomakkeessa on kysymyksessä olevana kautena merkitty kaikki arkipäivät a yhteen ajankäyttöryhmään, joka kuitenkin eri lomakkeissa vaihtelee). Tällöin on kysymyksessä tavallinen kvalitatiivinen analyysi, koska an jakautuu ominaisuuksilla j varustettuihin ryhmiin am_1 , am_2 jne. ja $\bar{p}_j = am_1/an$, am_2/an jne. $= m_1/n$, m_2/n jne. Siten tässä erikoistapauksessa

$$v_{x'_j}^2 = v_{\bar{x}_j}^2 = v_{\bar{p}_j}^2 = \bar{q}_j/n\bar{p}_j, \text{ jossa } \bar{q}_j = 1 - \bar{p}_j$$

(jos \bar{p}_j :llä on binomiaalinen todennäköisyysjakautuma).

Jos tarkastellaan x_{ij} -lukujen jakautumaa samojen edellytysten vallitessa, havaitaan, että ajankäyttöryhmässä j x_{ij} -luvut jakautuvat kahteen ryhmään siten, että ne $\bar{p}_j n$ koehenkilöllä ovat suuruudeltaan a ja $\bar{q}_j n$ koehenkilöllä $= 0$. Lukujen x_{ij} keskihajonta $s_{x_{ij}}$ on tässä tapauksessa suurin mahdollinen (edellyttäen, että \bar{p}_j pysyy samana). Tästä voidaan edelleen päätellä, että jos viimeksi esitetty kaava pannaan yleisempään muotoon

$$v_{\bar{p}_j}^2 = c_j^2 \bar{q}_j/n\bar{p}_j, \quad (9)$$

tämä kaava c_j -kertoimen arvolla 1 antaa keskivirheitä, joita todelliset keskivirheet eivät ylitä, vaikka x_{ij} ei olekaan $= a$.

Todellisuudessa ei päätutkimuksessa myöskään vaatimus $z_i = a$ tarkalleen toteudu. T-lomakkeissa oli kuitenkin vain verrattain harvoissa poikkeustapauksissa koko vuoden z_i -luku suurempi kuin 301 miespäivää, silloinkin se hyvin harvoin ylitti 310:tä. Pienempi kuin 301 se käsittelyohjeiden mukaan ei voinut olla työkykyisillä T-miehillä. Viimeksi mainittujen miesten koko vuoden z_i -lukujen keskiarvo koko näytteessä oli 301.5 miespäivää. Kun kuukausienkin z_i -luvut poikkeavat keskiarvostaan suhteellisesti yhtä vähän, voidaan sekä koko vuoden että kuukausien z_i -lukuja pitää käytännöllisesti katsoen vakioina ($= a$).

Näin ollen kaava 9 c_j -kertoimella 1 ilmaisee käytäntöön riittävällä varmuudella x'_j - ja \bar{p}_j -estimaattien suhteellisen varianssin ylärajan. Jos se pannaan muotoon

$$c_j^2 = s_{p_{ij}}^2 / \bar{p}_j \bar{q}_j \quad (10)$$

näkyä, että jonkin ajankäyttöryhmän c_j -kertoimen suuruus on riippuvainen vain kyseisen ryhmän p_{ij} - (ja samalla x_{ij} -)muuttujan keskihajonnasta ja keskiarvosta, ts. p_{ij} - (tai x_{ij} -)muuttujien prosenttisen frekvenssijakautumakuvaajan muodosta.

Kertoimen c_j riippuvaisuutta ajankäyttöryhmän j laadusta tutkittakoon ensin läänien ja koko maaseudun kaikkien työkykyisten koehenkilöiden näytteistä, erikseen koko vuoden ja kuukausien p_{ij} -luvuista. Jätetään toistaiseksi huomioon ottamatta arvion m' varianssi ja tarkastellaan vain m -ositteen sisäistä satunnaisvaihtelua.

Taulukkoon 19/s. 102 on laskettu kaavasta 1/s. 37 $s_{\bar{p}_j}$ -arviot 66:lle koko vuoden tunnusluvulle \bar{p}_j . Laskelmissa käytettiin m_1 -näytteinä eri läänien ja koko maaseudun kaikkia työkykyisiä koehenkilöitä ja m_2 -lukuna kaikissa tapauksissa koko maaseudun työkykyisten koehenkilöiden lukumäärää 44 667 miestä. Tulokset näkyvät myös kuvassa 6/s. 105, johon ne on merkitty koordinaatistoon \bar{p}_j -lukujen funktioina (ajankäyttöryhmistä I—VII ympyröillä ja ryhmistä VIII—XII pisteillä).

Saaduista $s_{\bar{p}_j}$ - ja niitä vastaavista \bar{p}_j -estimaateista on taulukkoon laskettu c_j -kerroin seuraavasti:

$$c_j^2 = m s_{\bar{p}_j}^2 / \bar{p}_j \bar{q}_j,$$

jossa $m = 44\,667$.

Kertoimille c_j laskettiin tämän jälkeen suoraviivaiset tasoitusarvot c_d ja c_t minimoimalla neliösummat

$$\sum_{r=1}^g \left(s_{\bar{p}_r}^2 - \frac{c_d^2}{m} \bar{p}_r \bar{q}_r \right)^2 \text{ ja}$$

$$\sum_{r=1}^g \left(\bar{p}_r \bar{q}_r - \frac{m}{c_t^2} s_{\bar{p}_r}^2 \right)^2, \text{ joissa}$$

r tarkoittaa $(\bar{p}, s_{\bar{p}})$ -parin järjestysnumeroa ja g mainittujen lukuparien lukumäärää. Selvyyden vuoksi on jätetty merkitsemättä alaindeksi j . Merkitsemällä neliösummien derivaatat 0:ksi, saadaan

Taulukko 19. Kesquivirhekaavan $s_{\bar{p}_j} = c_j \sqrt{\bar{p}_j \bar{q}_j / m}$ tunnusluvut \bar{p}_j , $s_{\bar{p}_j}$ ja c_j eräissä ajankäyttöryhmissä. Työkykyiset miehet, päätutkimus, $m = 44667$.

Table 19. Estimates \bar{p}_j , $s_{\bar{p}_j}$ and c_j in standard error formula $s_{\bar{p}_j} = c_j \sqrt{\bar{p}_j \bar{q}_j / m}$ for some activity groups. All able males, Main Survey, $m = 44667$.

Ajankäyttöryhmä (Ryhmän numero T-lomakkeessa) Activity group	Alue District	\bar{p}_j %	$s_{\bar{p}_j}$ %	c_j
1. Ansiotyö (08—19, 28) Paid work, total	Uudenmaan l. — Uusimaa C.	56.1	0.196	0.83
	Turun ja Porin l. — Turku and Pori C.	46.1	0.189	0.80
	Hämeen l. — Häme C.	50.9	0.183	0.77
	Kymen l. — Kymi C.	50.0	0.190	0.80
	Mikkelin l. — Mikkelä C.	42.3	0.173	0.74
	Kuopion l. — Kuopio C.	44.8	0.170	0.72
	Vaasan l. — Vaasa C.	39.4	0.177	0.77
	Oulun l. — Oulu C.	43.7	0.162	0.69
	Lapin l. — Lapland C.	45.2	0.154	0.65
Koko maaseutu — All rural districts	45.6	0.178	0.76	
2. Talonrakennus, ansiotyö (15, 16) House construction, paid work	Hämeen l. — Häme C.	8.0	0.094	0.73
3. Maatilatalous, metsätalous ja uitto (08—12, 14, 17, 22—27) Farming, forestry and floating, total	Turun ja Porin l. — Turku and Pori C.	48.6	0.194	0.82
	Koko maaseutu — All rural districts	42.8	0.158	0.67
4. Maatilatalous (08, 14, 17, 22, 23, 25—27) Farming, total	Koko maaseutu — All rural districts	42.8	0.158	0.67
5. Maatilatalous, oma työ (22—27) Farming, unpaid work	Uudenmaan l. — Uusimaa C.	26.8	0.169	0.81
	Turun ja Porin l. — Turku and Pori C.	36.4	0.178	0.78
	Ahvenanmaa — Aaland	34.4	0.165	0.73
	Hämeen l. — Häme C.	32.0	0.179	0.81
	Kymen l. — Kymi C.	33.3	0.180	0.81
	Mikkelin l. — Mikkelä C.	40.2	0.173	0.75
	Kuopion l. — Kuopio C.	37.3	0.164	0.72
	Vaasan l. — Vaasa C.	41.9	0.172	0.74
	Oulun l. — Oulu C.	39.3	0.157	0.68
	Lapin l. — Lapland C.	33.7	0.145	0.65
Koko maaseutu — All rural districts	36.5	0.172	0.75	
6. Maatalous (14, 17, 26, 27) Agriculture, total	Turun ja Porin l. — Turku and Pori C.	34.8	0.146	0.65
7. Maatalous, oma työ (26, 27) Agriculture, unpaid work	Turun ja Porin l. — Turku and Pori C.	27.8	0.139	0.66

Taulukko 19 (jatkoa) — Table 19 (continued)

8. Metsätalous ja uitto (08—12, 22—24) Forestry and floating, total	Uudenmaan l. — Uusimaa C.	7.8	0.056	0.44
	Turun ja Porin l. — Turku and Pori C.	10.8	0.065	0.44
	Ahvenanmaa — Aaland	7.8	0.043	0.34
	Hämeen l. — Häme C.	14.1	0.085	0.52
	Kymen l. — Kymi C.	13.3	0.084	0.52
	Mikkelin l. — Mikkelä C.	23.9	0.096	0.48
	Kuopion l. — Kuopio C.	21.5	0.099	0.51
	Vaasan l. — Vaasa C.	15.6	0.080	0.47
	Oulun l. — Oulu C.	19.7	0.095	0.50
Lapin l. — Lapland C.	23.4	0.105	0.52	
Koko maaseutu — All rural districts	16.2	0.084	0.48	
9. Metsätalous ja uitto, ansiotyö (08—12) Forestry and floating, paid work	Uudenmaan l. — Uusimaa C.	3.9	0.046	0.50
	Turun ja Porin l. — Turku and Pori C.	5.1	0.053	0.51
	Hämeen l. — Häme C.	8.8	0.079	0.59
	Kymen l. — Kymi C.	6.6	0.074	0.63
	Mikkelin l. — Mikkelä C.	15.0	0.120	0.71
	Kuopion l. — Kuopio C.	13.7	0.096	0.59
	Vaasan l. — Vaasa C.	7.7	0.071	0.56
	Oulun l. — Oulu C.	11.5	0.085	0.56
	Lapin l. — Lapland C.	15.5	0.096	0.56
Koko maaseutu — All rural districts	9.3	0.078	0.57	
10. Markkinapuun hakkuu, ansiotyö (09) Felling of wood for sale, paid work	Uudenmaan l. — Uusimaa C.	1.9	0.034	0.53
	Turun ja Porin l. — Turku and Pori C.	2.4	0.038	0.52
	Hämeen l. — Häme C.	4.5	0.057	0.58
	Kymen l. — Kymi C.	3.3	0.052	0.62
	Mikkelin l. — Mikkelä C.	7.9	0.074	0.58
	Kuopion l. — Kuopio C.	7.3	0.069	0.56
	Vaasan l. — Vaasa C.	3.9	0.052	0.57
Oulun l. — Oulu C.	5.3	0.061	0.58	
Lapin l. — Lapland C.	6.1	0.062	0.55	
Koko maaseutu — All rural districts	4.6	0.057	0.58	
11. Markkinapuun ajo vieraan metsästä (10) Hauling of wood for sale, paid work	Turun ja Porin l. — Turku and Pori C.	1.0	0.019	0.40
	Hämeen l. — Häme C.	1.9	0.029	0.45
	Mikkelin l. — Mikkelä C.	3.8	0.043	0.48
	Kuopion l. — Kuopio C.	2.7	0.036	0.47
	Vaasan l. — Vaasa C.	1.6	0.026	0.44
	Oulun l. — Oulu C.	2.4	0.033	0.46
	Lapin l. — Lapland C.	3.6	0.042	0.48
Koko maaseutu — All rural districts	2.0	0.030	0.45	
12. Työttömyysaika (05, 30, 31) Unemployment, underemployment and relief work	Koko maaseutu — All rural districts	4.0	0.045	0.49

Pienimmän neliösumman menetelmällä laskettu c-kerroin ryhmille $\begin{cases} 1-7: c_t = 0.75 \\ 8-12: c_t = 0.55 \end{cases}$
The coefficient c computed by the least squares method, for groups $\begin{cases} 1-12: c_t = 0.73 \end{cases}$

$$c_d^2 = \frac{m \sum_{r=1}^g s_{\bar{p}_r}^2 \bar{p}_r \bar{q}_r}{\sum_{r=1}^g (\bar{p}_r \bar{q}_r)^2} \quad (11)$$

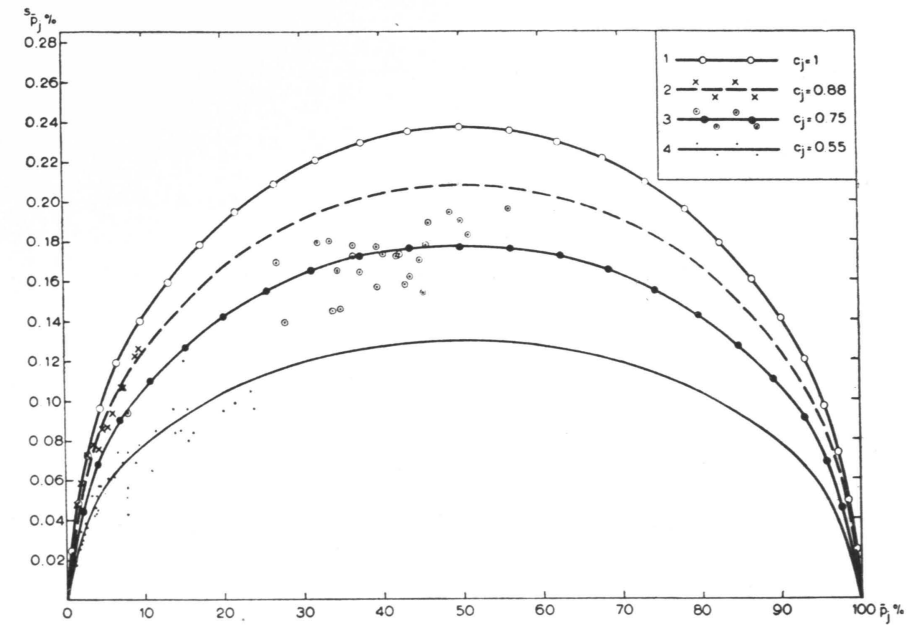
$$c_t^2 = \frac{m \sum_{r=1}^g s_{\bar{p}_r}^4}{\sum_{r=1}^g s_{\bar{p}_r}^2 \bar{p}_r \bar{q}_r} \quad (11 a)$$

Tasoitusarvot c_d ja c_t ovat origon kautta kulkevien regressiosuorien kulmakertoimia, joista $c_t > c_d$. Kun c_t sijoitettuna kaavaan 9/s. 100 c_j :n paikalle antaa suuremman keskivirheen kuin c_d , käytettiin seuraavissa laskutoimituksissa varovaisuuden vuoksi kaavaa 11 a ja vain vertailun vuoksi laskettiin myös eräitä c_d -estimaatteja. Huomattakoon, että koordinaatistoon $y = ms_{\bar{p}_j}^2 = s_{p_{ij}}^2$ ja $x = \bar{p}_j \bar{q}_j$ piirretyt varianssipisteet vaikuttavat regressiosuorien kulmakertoimiin sitä suuremmalla painolla, mitä etäämmällä ne ovat origosta (vrt. DEMING 1948, ss. 32—33). Kun aineistosta voitiin laskea suhteellisen vähän variansseja \bar{p}_j -estimaateille $> 20\%$, tällainen c_j -estimaattien tasoitusarvojen laskentatapa on perusteltavissa.

Kaikista taulukkoon 19/s. 102 merkityistä $(\bar{p}_j, s_{\bar{p}_j})$ -lukupareista saatiin $c_t = 0.78$, $c_d = 0.71$. Ajankäyttöryhmät 1—7 ($c_t = 0.75 = c_d$) ja 8—12 ($c_t = 0.55$, $c_d = 0.58$) eroavat selvästi toisistaan c_j -kertoimen keskimääräisen suuruuden perusteella. Yhtälön $s_{\bar{p}_j} = c_t \sqrt{\bar{p}_j \bar{q}_j / m}$ kuvaajat on piirretty kuvaan 6/s. 105 c_t -kertoimilla 1, 0.75 ja 0.55. Kuvaa tarkasteltaessa on huomattava, että $s_{\bar{p}_j} = s_{\bar{q}_j}$, vaikka viimeksi mainittuja lukuja selvyden vuoksi ei olekaan merkitty koordinaatistoon. Kuvasta näkyy, että tasoitusyhtälöt kertoimilla 0.75 ja 0.55 antavat asianomaisille \bar{p}_j -luville muutamia poikkeuksia lukuun ottamatta verrattain oikeat keskivirheet.

Kerroin c_j näyttää olevan hyvin vähän riippuvainen siitä, onko kysymyksessä koko maaseudun vai läänin näyte (poikkeuksena on Ahvenanmaa). Tästä voidaan päätellä, että c_j -kertoimet ovat samaa suuruusluokkaa myös työvoimapiirien ja metsänhoitolautakuntien toimialueiden näytteissä.

Kuukausien \bar{p}_j -estimaateille oli reikäkorttiaineistojen avulla teknillisesti vielä työläämpää laskea kaavasta 1/s. 37 keskivirheitä kuin koko vuo-



Kuva 6. Päättökimpuksen kaikkien työkykyisten miesten eräiden \bar{p}_j -estimaattien keskivirheet ($s_{\bar{p}_j}$) sekä niille lasketun tasoitusyhtälön $s_{\bar{p}_j} = c_t \sqrt{\bar{p}_j \bar{q}_j / m}$ kuvaajat. $m = 44667$. 2) Kuukausien estimaateille taulukosta 20/s. 105. 3) Koko vuoden estimaateille taulukon 19/s. 102 ajankäyttöryhmistä 1—7. 4) Sama ajankäyttöryhmistä 8—12. Vertailun vuoksi on kuvaan piirretty myös maksimikeskivirheiden kuvaaja 1).

Figure 6. Standard errors ($s_{\bar{p}_j}$) for selected \bar{p}_j estimates for all able males in the Main Survey and the corresponding curves for the empirical equation $s_{\bar{p}_j} = c_t \sqrt{\bar{p}_j \bar{q}_j / m}$ where $m = 44667$. 2) Estimates for selected months from Table 20/p. 106. For whole year from Table 19/p. 102, activity groups 1—7. 4) The same, activity groups 8—12. For comparison, the maximum standard error curve is also drawn in the figure.

den tunnusluville. Taulukossa 20/s. 106 esitetään vain kahden metsätyöpanosryhmän keskivirheet ja c_j -kertoimet koko maaseudulle eräinä kuukausina. Tulos osoittaa, että c_j -kerroin kaikissa tutkituissa tapauksissa on suhteellisen suuri koko vuoden c_j -kertoimiin verrattuna. Kuukausien c_j -kertoimet poikkeavat varsin vähän tasoitusarvostaan $c_t = 0.88$. Myös taulukon 20 keskivirheet on merkitty kuvaan 6 (rasteilla), samoin niiden tasoitusyhtälön kuvaaja.

Taulukko 20. Keskivirhekaavan $s_{\bar{p}_j} = c_j \sqrt{\bar{p}_j \bar{q}_j / m}$ tunnusluvut \bar{p}_j , $s_{\bar{p}_j}$ ja c_j kahdessa metsätyöpanosryhmässä eräinä kuukausina ja v. 1950. Koko maaseutu, työkykyiset miehet, päätutkimus, $m = 44\ 667$.

Table 20. Estimates \bar{p}_j , $s_{\bar{p}_j}$ and c_j of the standard error formula $s_{\bar{p}_j} = c_j \sqrt{\bar{p}_j \bar{q}_j / m}$ for two forest labour input groups. Monthly values and total year 1950. All rural districts and all able males. Main Survey, $m = 44667$.

Ajankäyttöryhmä (Ryhmän numero T-lomakkeessa) Activity group	Kuukausi — Month								Vuosi Year 1950	
	I	II	III	IV	VI	VIII	X	XII		
Markkinapuun hakkuu, ansio- työ (09) — <i>Felling of wood</i> for sale, paid work	\bar{p}_j %	8.9	9.4	7.2	4.7	2.0	1.4	3.6	7.3	6.1
	$s_{\bar{p}_j}$ %	0.122	0.126	0.107	0.086	0.059	0.048	0.078	0.107	0.057
	c_j	0.91	0.91	0.87	0.86	0.89	0.86	0.89	0.87	0.58
Markkinapuun ajo, ansiotyö (10) — <i>Hauling of wood for</i> sale, paid work	\bar{p}_j %	5.3	6.1	4.3	2.9	2.0
	$s_{\bar{p}_j}$ %	0.087	0.094	0.076	0.073	0.030
	c_j	0.82	0.83	0.79	0.92	0.45

Koko vuoden kertoimen c_j ja sen tasoitusarvon c_t riippuvaisuutta ajankäyttöryhmän laadusta tutkittiin myös pienalueittain, nimittäin päätutkimuksen Ranuan kunnan näytteestä taulukon 15/s. 86 perusteella sekä II esitutkimuksen Ranuan ja Kontiolahden kuntien yhteisestä näytteestä taulukon 3/s. 39 perusteella.

Taulukkoon 15 merkityissä H-tutkimuksen (Ranuan kunta) ajankäyttöryhmissä oli c_j -kerroin yleensä suuruusluokkaa 0.3 ja $c_t = 0.41$. Yli 0.50 oli c_j -kerroin seuraavissa ryhmissä (ryhmien numerot tarkoittavat vastaavia taulukon 15 ryhmän numeroita, ryhmien nimet osoittavat niiden pääasiallisen sisällön):

	c_j
2. Uitto- ja tietyöpanos	0.52
4. Kaupunkielinkeinojen työpanos	0.69
1—4. Ansiotyöpanos	0.56
5—9. Oman työn panos	0.51
10. Vapaa-aika	0.69

Taulukkoon 3/s. 39 merkityissä II esitutkimuksen (Ranuan ja Kontiolahden kunnat) ajankäyttöryhmissä c_j -kertoimet vaihtelivat rajoissa 0.24 — 0.94 ja $c_t = 0.68$. Yli 0.50 oli c_j -kerroin seuraavissa ryhmissä (ryhmissä numerot tarkoittavat taulukon 3 vastaavia ajankäyttöryhmien numeroita, ryhmien nimet ilmaisevat niiden pääasiallisen sisällön):

	c_j
1. Metsänhakuun ansiotyöpanos	0.56
4. Uittotyöpanos	0.75
6. Tietyöpanos	0.52
7. Saha- ja lastaustyöpanos	0.51
8. Talonrakennuksen ansiotyöpanos	0.67
9. Kaupunkielinkeinojen työpanos	0.86
10. Julkisten tehtävien, opiskelu-, asevelvollisuus aika	0.59
16. Kotiteollisuustyöpanos	0.71
17. Poronhoitotyöpanos	0.94
18. Maatalouden oman työn panos	0.51
26. Ansiotyöpanos	0.71
27. Oman työn panos	0.69

Myös kuukausittain laskettiin päätutkimuksen Ranuan kunnan näytteestä (työkykyiset miehet) c_j -kerroin eräissä yhdistetyissä ajankäyttöryhmissä. Tulokset näkyvät taulukosta 21/s. 108. Kuukausien c_j -kertoimet ovat nytkin suuremmat kuin koko vuoden vastaava c_j -kerroin. Maatalouden oman työn panos on poikkeus sikäli, että kuukauden c_j -kerroin eräissä tapauksissa on jopa pienempi kuin vastaava koko vuoden c_j . Tämä johtuu ilmeisesti siitä, että syyskuukaudet ovat siirtymäaikaa, jolloin kuukauden kokonaistyöpanos monilla miehillä jakautuu useiden työpanosryhmien kesken.

Tähänastisessa c_j -kertoimen tarkastelussa on ollut kysymys jonkin alueen *kaikki työkykyiset miehet* käsittävästä näytteen osasta. Päätutkimuksen tuloksissa esitetään x'_j - ja \bar{p}_j -estimaatteja myös pienemmille näytteen alaryhmille, esimerkiksi metsätyöllisille ja kääpiötilojen viljelijöille (viljelijällä tarkoitetaan työkykyistä miestä, jonka maatalouden oman työn panos v. 1950 oli > 20 päivää. Kääpiötilan viljelijä on viljelijä, joka 31. 12. 1950 asui viljelmällä, jossa oli peltoa 0.25 — 4.99 hehtaaria). Läänien näytteistä laskettiin seuraavat *kääpiötilojen viljelijöiden* koko vuoden c_j -kertoimet ($c_t = 0.5$).

Taulukko 21. Keskivirhekaavan $s_{\bar{p}_j} = c_j \sqrt{\bar{p}_j \bar{q}_j / m}$ kerroin c_j eräissä yhdistetyissä ajankäyttöryhmissä eräinä kuukausina ja v. 1950. Ranua, työkykyiset miehet, päätutkimus.

Table 21. Coefficient c_j of the standard error formula $s_{\bar{p}_j} = c_j \sqrt{\bar{p}_j \bar{q}_j / m}$ for some combined activity groups, monthly values and total year 1950. Ranua, all able males, Main Survey.

Ajankäyttöryhmä (Ryhmän numero T-lomakkeessa) Activity group	Kuukausi — Month						Vuosi Year 1950
	I	III	V	VII	IX	XI	
Metsätyöpanos vieraalle (08—11) — Input of paid labour to forestry	0.95	0.98	0.64	0.71	0.80	0.53	0.40
Maatalouden oman työn panos (27) Input of unpaid labour to agriculture	0.73	0.62	0.52	0.52	0.32	0.37	0.51
Muu arki aika (01—07) — Non-working week-days	0.55	0.61	0.62	0.59	0.66	0.75	0.37

	c_j
Maatilatalouden oman työn panos, Turun ja Porin lääni	0.56
Kuopion lääni	0.50
Mikkelin lääni	0.53
Maatalouden oman työn panos, Lapin lääni	0.44
Metsätalouden ja uiton työpanos, Oulun lääni	0.50
Mikkelin lääni	0.49
Talonrakennuksen työpanos, Vaasan lääni	0.60
Lapin lääni	0.59
Ansiotyöpanos, Uudenmaan lääni	0.66
Työttömyysaika, Lapin lääni	0.49

Vain talonrakennus- ja koko ansiotyöpanoksessa saatiin paljon suurempi c_j -kerroin kuin 0.5. On odotettavissa, että kääpiöviljelijöillä oman maatilatalouden työpanosryhmissä c_j -kerroin on huomattavasti pienempi kuin kaikilla työkykyisillä miehillä, koska kyseisen työvoiman ryhmä rajoitettiin niin, ettei oman maatilatalouden p_{ij} -lukuissa voi esiintyä pienempiä lukuja kuin 0.07.

Edellä selostetut c_j -kerrointa laskettaessa saadut tulokset osoittavat, että c_j vaihtelee varsin laajoissa rajoissa. Paitsi ajankäyttöryhmän laa-

dusta, se on riippuvainen myös siitä, kuinka suuresta alueesta on kysymys. Kuukausien p_{ij} -jakautumien c_j -kerroin on yleensä suurempi kuin koko vuoden vastaavissa jakautumissa. Lisäksi c_j on riippuvainen siitä, millä tavalla rajoitetusta, jonkin alueen näytteen osasta se lasketaan. On ilmeistä, ettei voida laskea harvalukuista määrää tasoitettuja c_t -kerroimia, jotka kaavassa $s_{\bar{p}_j} = c_t \sqrt{\bar{p}_j \bar{q}_j / m}$ antaisivat kaikille ajateltavissa oleville tunnusluvuille \bar{p}_j keskivirheen edes 10 %:n varmuudella.

Kertoimen c_t maksimiarvolla 1 saadaan hyvin lukuisissa tapauksissa 2-, jopa 4-kertainen keskivirhe. Jotta välttyttäisiin tältä käytännönkin tarkoituksiin kovin karkealta menettelyltä, tarkastellaan seuraavassa, mitkä seikat määräävät p_{ij} -lukuun jakautumakuvaajan tyyppin ja miten se puolestaan vaikuttaa c_j -kertoimeen.

Kun tätä jakautumakuvaajan muotoa kuvaillaan termeillä M-, A-, J-jakautuma jne., tarkoitetaan p_{ij} -lukuun jakautuman aritmeettisella asteikolla varustettuun koordinaatistoon siten piirrettyä kuvaajaa, että x-akselilla ovat p_{ij} -luvut 0:sta 1:een ja y-akselilla p_{ij} -lukuun suhteellinen lukuisuus. Teoreettisesti on laskettavissa c_j eri jakautumatyypeille seuraavasta professori L. TÖRNQVISTIN antamasta kaavasta

$$c_j^2 = \frac{\int_{x=0}^1 (x - \bar{p}_j)^2 dF_j(x)}{\bar{p}_j \bar{q}_j} \quad (10a)$$

Seuraavat toereettiset c_j -kertoimet laskettiin:

	\bar{p}_j	c_j
A- jakautumassa	1/2	$\sqrt{1/6} = 0.41$
Kolmiojakautumassa	1/3 ja 2/3	$\sqrt{1/4} = 0.50$
Tasaisessa jakautumassa	1/2	$\sqrt{1/3} = 0.58$
M- jakautumassa	1/2	$\sqrt{1/2} = 0.71$

Kuvaan 7/s. 111 on taulukon 22/s. 110 perusteella piirretty eräitä kokemusperäisiä, erilaisia c_j -kertoimia antavia jakautumia. Kuvassa on Y-akselin asteikko tosin logaritminen, koska kuvaajaa on vaikea selvästi esittää aritmeettisessä asteikossa. Tässä tapauksessa kuvaajien perustyyppi olisi kuitenkin sama myös aritmeettisessä asteikossa kuvaajaa 3 lukuun ottamatta, joka siinä olisi tyyppiä J sen sijaan, että se kuvassa on lähinnä kolmiojakautuman tyyppiä.

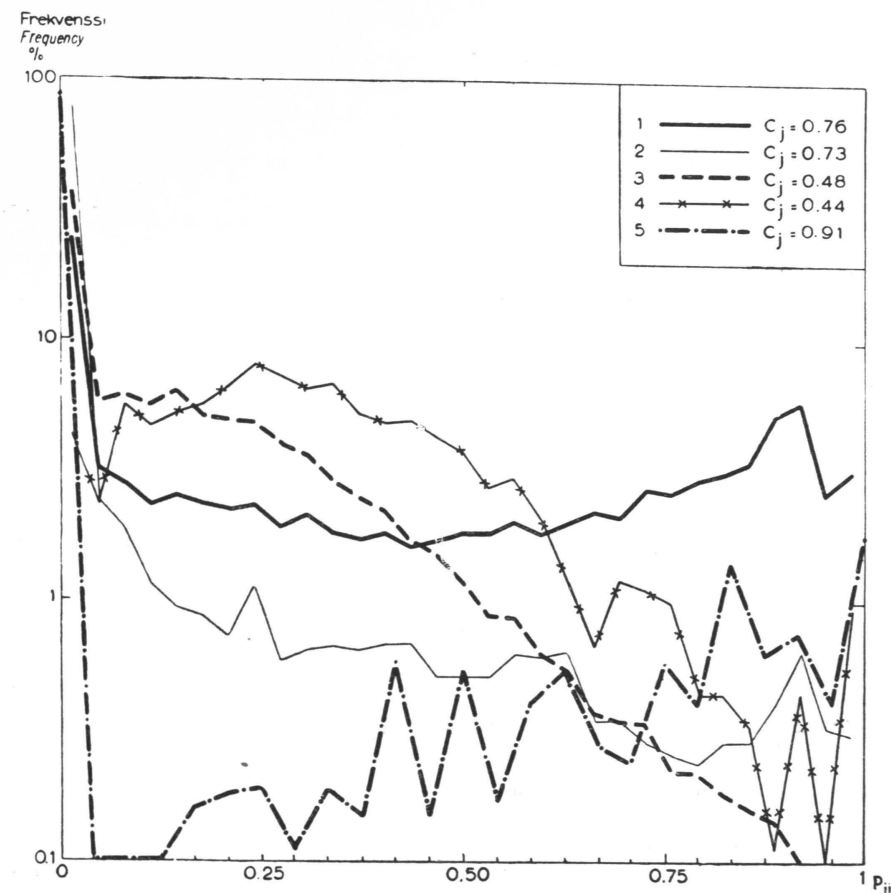
Taulukko 22. Eräiden ajankäyttöryhmien p_{ij} -muuttujan prosenttiset frekvenssijakautumat. Päätutkimus.

Table 22. Percentual frequency distribution of the variable p_{ij} in some activity groups. Main Survey.

p_{ij} %	Frekvenssi * — Frequency *				p_{ij} %	Frekvenssi * Frequency *
	1	2	3	4		
0.0— 3.3	24.4	80.0	37.0	4.3	0 tai or 4	89.9
3.4— 6.7	3.2	2.4	5.7	2.3	8	0.1
6.8— 10.0	2.8	1.9	6.1	5.6	13	0.1
10.1— 13.3	2.3	1.1	5.6	4.6	17	0.2
13.4— 16.6	2.5	0.9	6.3	5.1	21	0.2
16.7— 20.1	2.3	0.9	5.0	5.6	25	0.2
20.2— 23.4	2.1	0.7	4.9	6.3	29	0.1
23.5— 26.7	2.3	1.2	4.8	8.0	33	0.2
26.8— 30.0	1.9	0.6	4.0	7.1	38	0.1
30.1— 33.4	2.1	0.6	3.6	6.4	42	0.6
33.5— 36.7	1.8	0.7	2.9	6.7	46	0.2
36.8— 40.0	1.7	0.6	2.5	5.0	50	0.5
40.1— 43.3	1.8	0.7	2.2	4.8	54	0.2
43.4— 46.6	1.6	0.7	1.7	4.9	58	0.4
46.7— 50.0	1.7	0.5	1.4	4.0	63	0.5
50.1— 53.3	1.8	0.5	1.2	3.7	67	0.3
53.4— 56.6	1.8	0.5	0.9	2.7	71	0.2
56.7— 59.9	2.0	0.6	0.9	3.0	75	0.6
60.0— 63.3	1.8	0.6	0.6	2.1	79	0.4
63.4— 66.6	2.0	0.7	0.5	1.2	83	1.4
66.7— 69.9	2.2	0.3	0.4	0.7	88	0.6
70.0— 73.2	2.1	0.3	0.4	1.2	92	0.8
73.3— 76.5	2.7	0.3	0.3	1.1	96	0.4
76.6— 79.9	2.6	0.3	0.2	1.0	100	1.8
80.0— 83.2	2.9	0.2	0.2	0.4	.	.
83.3— 86.5	3.1	0.3	0.2	0.4	.	.
86.6— 89.8	3.4	0.3	0.2	0.3	.	.
89.9— 93.2	5.2	0.4	0.1	0.1	.	.
93.3— 96.6	5.9	0.6	0.1	0.4	.	.
96.7— 100.0	6.0	0.6	0.1	1.0	.	.
Yhteensä — Total	100.0	100.0	100.0	100.0	Yhteensä Total	100.0
m	44 365	5 086	44 667	916	m	44 667
c_j	0.76	0.73	0.48	0.44	c_j	0.91

- * 1. Ansiotyöpanos, koko maaseutu ilman Ahvenanmaata.
2. Talonrakennuksen ansiotyöpanos, Hämeen lääni.
3. Metsä- ja uittotyöpanos, koko maaseutu.
4. Maatalouden oman työn panos, Lapin lääni, kääpiötilojen viljelijät.
5. Markkinapuun hakkuun ansiotyöpanos, koko maaseutu, helmikuu 1950.

- * 1. Total input of paid labour, all rural districts, excl. Aaland.
2. Input of paid labour to house construction, Häme County.
3. Labour input to forestry and floating, all rural districts.
4. Input of unpaid labour to agriculture, Lapland County, farmers of dwarf holdings.
5. Input of paid labour to felling and hauling of timber for sale, all rural districts, February 1950.



Kuva 7. Eräiden ajankäyttöryhmien p_{ij} -muuttujan prosenttiset frekvenssikäyrät. Päätutkimus. Lähde: taulukko 22/s. 110. Ryhmän numeroiden 1—5 selitykset taulukon alaviitassa. Puolilogaritmisen asteikko.

Figure 7. Percentage frequency distribution curves of p_{ij} variables in selected activity groups. Main survey. Source: Table 22/p. 110. Group numbers 1—5 refer to the foot-note in the table. Semi-logarithmic scale.

Jakautumakuvaajan muodon ja c_j -kertoimen (yksinkertaisuuden vuoksi jätetään seuraavassa c :n alaindeksit pois, milloin epäselvyyttä ei voi syntyä) riippuvaisuussuhteesta voidaan edellä esitetyn perusteella päätellä seuraavaa.

1) Kun kuvaaja on tyyppiä M tai U siten, että molemmissa äärimäisissä luokissa on suhteellisen runsaasti p_{ij} -muuttujia, on $c > 0.55$ ja

sitä suurempi mitä pienempi osa p_{ij} -luvuista on ääriarvojen 0 ja 1 välillä ääriarvoja mukaan lukematta. Jos tämä osa on 0 %, on $c = 1$;

2) Kun kuvaaja on kolmio- tai J-jakautuman tyyppiä siten, että p_{ij} -lukuja toisessa ääriluokista on suhteellisesti hyvin vähän, on c hyvin harvoin > 0.55 ja lähenee lukua 0, kun väliarvojen osuus vähenee.

3) Kun molemmissa äärimmäisissä luokissa keskellä oleviin luokkiin verrattuna on suhteellisen vähän p_{ij} -muuttujia ja kuvaaja on lähimain katkaistun normaalikäyrän tai A:n muotoinen, on c myöskin < 0.55 ja tulee jälleen 0:ksi, jos kaikki muuttujat keskittyvät samaan luokkaan.

Edellä selostetun ja päätutkimuksen taulukossa 17/s. 97 esitettyä tyyppiä olevien tulosten perusteella voidaan tehdä seuraavat yleiset päätelmät:

1) *Kaikkien työkykyisten miesten koko vuoden 1950 arkiajan jakautumat.*

Koko maaseudun, läänien, työvoimapiirien ja metsänhoitolautakuntien toiminta-alueiden T-lomakkeen alkuperäisten (yhdistämättömien) ajankäyttöryhmien p_{ij} -muuttujien jakautumat ovat yleensä J-tyyppisiä, koska valtaosalla koehenkilöistä ajankäyttöryhmää ei ollenkaan esiinny ($p_{ij} = 0$) ja hyvin harvoilla miehillä $p_{ij} \sim 1$. Jos näissä ryhmissä käytetään kerrointa $c = 0.55$, saadaan likimäärin oikea tai korkeintaan ehkä kaksinkertainen keskivirhe. Seuraavat T-lomakkeen alkuperäiset ajankäyttöryhmät muodostavat poikkeuksen (numerot viittaavat T-lomakkeeseen).

- 01. Opiskelu- ja asevelvollisuusajaksi.
- 15 ja 16. Talonrakennuksen ansiotyöpanos.
- 17. Maatilatalouden ansiotyöpanos.
- 19. Kaupunkielinkeinojen työpanos (ryhmät 19 ja 28 esiintyvät tuloksissa vain yhdistettyinä).
- 20. Käsityön ja pienteollisuuden oman työn panos.
- 21. Metsästyksen, kalastuksen ja poronhoidon oman työn panos.
- 27. Maatalouden oman työn panos (uudisraivausta lukuun ottamatta).

Näissäkin ryhmissä $p_{ij} = 0$ suurimmalla osalla kaikista työkykyisistä miehistä, mutta myös $p_{ij} \sim 1$ on suhteellisen tavallinen. Jakautuman kuvaaja on siten muotoa M tai U ja $c > 0.55$. Kun läänin suuruusluokkaa olevilla alueilla esiintyy suhteellisen runsaasti myös kaikkia ääriarvojen välisiä p_{ij} -muuttujia, voidaan katsoa kertoimen $c = 0.75$ näissä ryhmissä antavan likimäärin oikean keskivirheen.

T-lomakkeen alkuperäisten ajankäyttöryhmien yhdistelmissä syntyville jakautuman muodoille on mahdotonta luoda yleistä sääntöä. Todetakaan kuitenkin, että jos yhdistettyyn ryhmään sisältyy joko T-lomak-

keen ryhmä 19 tai 27 (mutta ei molempia), tämä määrää suhteellisen suuruutensa takia myös c -kertoimen. Niinpä $c = 0.75$ on käyttökelpoinen ryhmissä ansiotyöpanos ja oman työn panos. Jos sen sijaan sekä ryhmä 19 että ryhmä 27 sisältyvät yhdistettyyn ryhmään (esim. koko työpanos), saa jakautuma J-muodon ja $c = 0.55$ on siten käyttökelpoinen. Lisäksi todettakoon, että koko metsä- ja uittotyöpanos kaikkine mahdollisine alaryhmineen on muotoa J ja $c = 0.55$ antaa niissäkin tyydyttävän tuloksen.

Tuloksissa esitetään työkykyisten miesten arviot x'_j ja \bar{p}_j myös niin pienille kuntaryhmille, että ao. näytekokoisten lukumäärä on 100—200. Silloin käytetään vain yhdistettyjä ajankäyttöryhmiä

metsätalouden työpanos,
 maatalouden työpanos,
 metsä- ja maatalouden ynnä puutavaran kaukokuljetuksen työpanos,
 kaupunkielinkeinojen työpanos,
 talon ja teiden rakennuksen työpanos sekä
 työttömyysaika.

Edellä todettiin, että kunnan näytteessä voi koehenkilöiden ajankäyttö olla huomattavasti tasalaatuisempi kuin läänien ja koko maaseudun näytteissä. Kun mainituissa päätutkimuksen tuloksissa kunnat pyrittiin yhdistämään mahdollisimman tasalaatuisiin naapurikuntien ryhmiin, on todennäköistä, että useissa kuntaryhmissä mainittujen ajankäyttöryhmien c_j -kerroin on huomattavasti pienempi kuin edellä esitetiin. Toisaalta todettiin edellä selostetuissa Ranuan sekä Ranuan ja Kontiolahden yhdistettyjen näytteiden c_j -kertoimissa poikkeuksellisen suuria lukuja, jopa uittotyöpanoksessa ja vapaa-ajassa. Tämä johtuu siitä, että pienillä alueilla saattaa jokin niille erikoinen ilmiö, kuten Ristisaaren erottelulaitoksen merkitys Kontiolahdella, vaikuttaa p_{ij} -muuttujan jakautumaan. Näin ollen on varmintä laskea näiden pienalueidenkin tunnuslukujen keskivirheet koko maaseudulle ja läänin kokoisille alueille johdetuilla c -kertoimilla, jollei jokin erikoistapaus vaadi harkinnanvaraista poikkeamista niistä.

2) *Metsätyöllisten, viljelijöiden ja muiden työvoiman erikoisryhmien koko vuoden 1950 arkiajan jakautumat.* Koko maaseutu, läänit ja metsänhoitolautakuntien toimialueet.

Kaikkien työkykyisten miesten mainitunlaiset alaryhmät on tuloksissa useimmiten erotettu pääryhmästään siten, että jonkin työpanoksen vähimmäismäärä on rajoitettu. Niinpä metsätyöllisillä tuli metsä- ja uittotyö-

panoksen v. 1950 olla vähintään 1 päivä, viljelijöillä tuli maatilatalouden oman työn panoksen olla suurempi kuin 20 päivää. Tästä johtuu, että ainakin kyseisessä työpanosryhmässä tutkittavan joukon c_j -kerroin on pienempi kuin kaikkien työkykyisten miesten vastaavassa ryhmässä. Niinpä viljelijöillä voidaan käyttää niissäkin työpanosryhmissä, joihin maatalouden oman työn panos sisältyy, kerrointa $c = 0.55$. Metsätyöllisten metsä- ja uittotyöpanoksen sekä sen alaryhmien $c = 0.40$ antanee tyydyttävän tarkat keskivirheet.

Jos koko näytteestä erotetaan alaryhmä jonkin muun ominaisuuden kuin työpanoksen suuruuden perusteella, esim. määrätty ikäryhmä, ei ole syytä poiketa kaikkien työkykyisten miesten c -kertoimista.

3) *Kuukausien arkiajan jakautumat.* Kaikki työkykyiset miehet ja metsätyölliset. Koko maaseutu ja työvoimapiirit.

Kuukausien p_{ij} -muuttujat ovat hyvin usein vahvasti keskittyneet ääriarvoihin. Tosin muunkin tyyppisiä jakautumia voi esiintyä, mutta yleistä sääntöä niiden riippuvaisuudesta ei voida muodostaa. Jotta vältyttäisiin liian pieniltä keskivirheiltä, on kuukausien x'_j - ja \bar{p}_j -estimaatteja testattaessa paras käyttää kerrointa $c = 0.9$.

Määrittämällä c -kerroin edellä esitettyjen sääntöjen avulla saadaan — ottamalla huomioon myös m -ositteiden välinen varianssi — seuraavista kaavoista x'_j - ja \bar{p}_j -estimaattien varianssit käyttökelpoisella tarkkuudella. Kaavat on saatu noudattamalla samaa periaatetta kuin kaavoissa 6, 6 a ja 8/s. 96:

$$I \quad v_{x'_j}^2 = \frac{(1-\varphi)}{m} + \frac{c^2 (z' - x'_j)}{m x'_j} = \frac{(1-\varphi)}{m} + \frac{c^2 \bar{q}_j}{m \bar{p}_j} \quad (12)$$

Kun on kysymys kaikkien työkykyisten miesten tunnusluvuista, on $\frac{(1-\varphi)}{m}$ niin pieni, että se voidaan jättää huomioon ottamatta. Jos $\bar{p}_j \gtrsim 20\%$ voidaan POISSONIN todennäköisyysjakautuman mukaisesti jättää huomioon ottamatta — x'_j ja vastaavasti \bar{q}_j . Kun nämä molemmat vaatimukset ovat täytetyt ja lisäksi pidetään muuttujaa z_i vakiona $a = \frac{z'}{m'}$, on

$$II \quad v_{x'_j}^2 = c^2 Ka \frac{1}{x'_j} = \frac{c^2}{m \bar{p}_j} \quad (12 a)$$

Estimaatin x'_j absoluuttinen varianssi tulee vastaavasti olemaan likiarvoltaan

$$I \quad s_{x'_j}^2 = \frac{(1-\varphi)}{m} x'^2 + \frac{c^2}{m} (z' - x'_j) x'_j \quad (13)$$

Kaikilla työkykyisillä miehillä, kun $\bar{p}_j \gtrsim 20\%$, voidaan käyttää lyhennettyä kaavaa

$$II \quad s_{x'_j}^2 = c^2 Ka x'_j \quad (13 a)$$

Estimaatin \bar{p}_j -varienssi saadaan kaavasta

$$I \quad s_{\bar{p}_j}^2 = c^2 \bar{p}_j \bar{q}_j / m \quad (14)$$

ja kun $\bar{p}_j \gtrsim 20\%$ voidaan käyttää lyhennettyä kaavaa

$$II \quad s_{\bar{p}_j}^2 = c^2 \bar{p}_j / m \quad (14 a)$$

Asettamalla kaavaparissa 12 ja 12 a/s. 114 oikeat puolet yhtä suuriksi saadaan $c^2 = (1 - \varphi)$. Kun tämä ehto on likipitäen täytetty, voidaan siten supistettuja kaavoja 12 a ja 13 a käyttää muista kaavojen käyttöä varten annetuista säännöistä välittämättä.

Taulukkoon 18/s. 99 on laskettu keskivirheet $s_{x'_j}$ kaavoista 13 ja 13 a. Keskivirheet ovat myös suurissa ajankäyttöryhmissä likimäärin samat, mikä johtuu siitä, että tässä $(1 - \varphi) \sim 0.55^2$.

335. Tulokset toisista lähteistä saatujen tietojen valossa

3351. Viljelijäperheiden miestyövoiman käyttö

Päätutkimuksen tulosten vertailun toisista lähteistä saatuihin vastaviin tietoihin voi odottaa antavan lisävalaistusta käytetyn menetelmän luotettavuuteen. Vertailukohteiden valinta suoritettiin niin, että vuoden 1950 maaseudun työvoiman tutkimuksen (jota tässä lyhyiden vuoksi nimitetään A-tutkimukseksi) tuloksia verrattiin kussakin tapauksessa eri lähteistä saatuihin tietoihin.

Vertailukohteet olivat seuraavat (luvussa 335 käytettyine lyhennysmerkintöineen): viljelijäperheiden oman miestyövoiman käyttö v. 1950 maatalouslaskennan mukaan (B-tutkimus), metsä- ja uittotyöntekijöiden lukumäärä vuoden 1950 lopussa väestönlaskennan mukaan (C-tutkimus),

metsätöiden määrä ja kausivaihtelu v. 1950 kulkulaitosten ja yleisten töiden ministeriön metsätyövoimatilaston mukaan (D-tutkimus) sekä työttömyysajan määrä ja kausivaihtelu v. 1950 kuntien työttömyyskortistojen mukaan (E-tutkimus).

Maatalouslaskennan erikoisselvityksiä varten poimittiin vuoden 1950 lopussa valtakunnan 2 ha ja sitä enemmän peltoa käsittävistä maatiloista viljelmäsuuruusluokittain ositettu otos, josta tutkittiin mm. maatioilla v. 1950 käytetty (ihmis)työvoiman määrä ja sen koostumus (Yleinen maatalouslaskenta ... 1954). Samalla saatiin tiedot myös viljelijäperheen ansiotyöpanoksesta. Mittausmenetelmänä oli haastattelu, joka kuitenkin poikkesi A-tutkimuksen haastattelumenetelmästä seuraavissa olennaisissa kohdissa. Haastattelun kohteena oli tavallisesti maatalon isäntä, joka siten joutui antamaan tiedot paitsi omasta myös muiden viljelijäperheen jäsenten työpanoksesta sekä maatilalla käytetyn palkkatyövoiman työpanoksesta. Tiedot työpäivien määristä haastateltavan tuli ilmoittaa koko kalenterivuodelta sitä kausiin jakamatta. Koska maatalouslaskennassa kysyttiin vain osaa viljelijäperheen jäsenten ja työntekijöiden vuoden käytöstä, ei tällöin useinkaan ollut mahdollisuutta tarkistaa ilmoitettujen työpanosten summan paikkansapitävyyttä, kuten A-tutkimuksessa voitiin tehdä. B-tutkimuksen etu A-tutkimukseen verrattuna oli haastattelijoiden suurempi pätevyys (maatalousteknikoita) sekä vastausten kadon vähäisyys.

B-tutkimuksen tuloksista voidaan vain viljelijäperheeseen kuuluvien 15-vuotiaiden ja sitä vanhempien miesten työpanostietoja verrata A-tutkimuksen tuloksiin. A-tutkimuksen perusteella laskettiin 2—4.99 ha peltoa käsittävien maatilojen miespuolisten viljelijöiden arkiajan käyttö. *Viljelijöillä* tarkoitettiin sellaisia ao. maatioilla vuoden 1950 päättyessä asuneita 15—64-vuotiaita työkykyisiä tilan haltijaperheen jäseniä, jotka vuoden 1950 aikana olivat tehneet tilan maatalous-, metsätalous-, kotitalous- tai rakennustöitä yhteensä vähintään 21 työpäivää. Mainittuja A-tutkimuksesta saatuja tietoja on taulukossa 23 verrattu B-tutkimuksen 2—4.99 hehtaaria peltoa käsittävien maatilojen 14 vuotta vanhempien viljelijäperheeseen kuuluvien miesten työpanostietoihin, jotka laskettiin maataloushallituksesta saatujen lukujen perusteella.

B-tutkimuksen lukuihin sisältyy A-tutkimuksessa työkyvyttömiksi merkittyjen sekä 64 vuotta vanhempien miesten työpanos, kaupala- ja kaupunkikuntien ao. maatilojen työpanos sekä vielä niidenkin viljelijäperheeseen kuuluneiden miesten työpanos, joita edellä annetun määritelmän mukaan ei luettu A-tutkimuksessa viljelijöihin. Kaikki nämä työpanokset kuitenkin puuttuvat taulukkoon merkityistä A-tutkimuksen

Taulukko 23. Viljelijäperheeseen kuuluvien miesten arkiajan käyttö. Maatalon suuruusluokka 2—4.99 ha peltoa. Päättutkimuksen (A) ja v:n 1950 maatalouslaskennan otoksen (B) vertailu. Koko maaseutu. Varmuusvälit on merkitty kaksinkertaisin keskivirhein.

Table 23. Activity of males belonging to farmer families. Farm size-group 2—4.99 ha under plough. Comparison between Main Survey (A) and sample of 1950 Census of Agriculture (B). All rural districts. Confidence intervals with doubled standard errors.

Ajankäyttöryhmä Activity group	¹ A	² B	⁵ A/B %
	1 000 miespäivää — man-days in thousands		
1. Kotitarvepuun hakkuu ja ajo sekä metsänhoito, oma työ — <i>Felling and hauling of wood for domestic use, silviculture, unpaid work</i>	2 816 ± 148	2 596	108 ± 6
2. Markkinapuun hakkuu ja ajo, oma työ — <i>Felling and hauling of wood for sale, unpaid work</i>	³ 1 054 ± 86	⁴ 842	125 ± 10
3. Uudisrakennus, oma työ — <i>Construction of new houses, unpaid work</i> ..	1 319 ± 128	1 339	99 ± 10
4. Uudisraivaus, oma työ — <i>Land clearing, unpaid work</i>	1 408 ± 99	1 420	99 ± 7
5. Muu maatalous, oma työ. — <i>Other agricultural employment, unpaid work</i>	12 551 ± 364	16 968	74 ± 2
6. Yhteensä ryhmät 1—5 — <i>Total of groups 1—5</i>	19 148 ± 494	23 165	83 ± 2
7. Metsästys, kalastus, poronhoito, oma käsityö ja pienteollisuus — <i>Hunting, fishing, reindeer raising, domestic handicraft and minor industries</i>	1 097 ± 116
8. Avoin työttömyys ja hukka-aika — <i>Unemployment and wasted days</i>	1 193 ± 92
9. Yhteensä ryhmät 6—8 — <i>Total of groups 6—8</i>	21 438 ± 536	23 165	93 ± 2
10. Ansiotyöpanos. — <i>Input of paid labour</i> ..	9 871 ± 409	10 637	93 ± 4
11. Yhteensä ryhmät 9—10 — Total of groups 9—10	31 309 ± 726	33 802	93 ± 2

¹ Työkykyiset 15—64-vuotiaat miehet. — *Able males between 15—64 years.*

² Kaikki yli 14-vuotiaat miehet. — *All males 15 years and over.*

³ Myös pystyleimikon ostajan palkkaama työ. — *Incl. labour employed by the buyer of a parcel of timber on the stump.*

⁴ Vain työpanos hankintaleimikoissa. — *Incl. only labour input in stands marked for delivery cutting.*

⁵ Koska B-lukujen keskivirheitä ei tunneta, ei niitä ole voitu ottaa huomioon. Todellisuudessa ovat varmuusvälit siten sarakkeeseen merkityjä laajemmat. — *As the standard errors of the B investigation are not known they cannot be taken into account. Thus the true confidence intervals are larger than those marked in the column.*

lukuista. Näin muodoin tulisi viimeksi mainittujen olla jonkin verran pienempiä tai ei ainakaan suurempia kuin B-tutkimuksen luvut. Poikkeuksen muodostaa ryhmä »markkinapuun hakkuu ja ajo», johon A-tutkimuksessa sisällytettiin myös maatilalan metsässä puutavaran pystyyn ostajan palkkaamana suoritettut työt. B-tutkimuksessa näitä töitä ei luettu mainittuun ryhmään ja on epäiltävää, tulivatko ne myöskään ryhmään »ansiotyöt oman maatilalan ulkopuolella». Ero tässä ryhmässä ei siis johdu A-tutkimuksen harhaisuudesta.

Ryhmässä »kotitarvepuun hakkuu ja ajo sekä metsänhoitotyöt» voinee A- ja B-lukujen erotus olla merkitsevä. Tämä seikka saattaa johtua siitä, että A-tutkimuksessa haastattelijoiden puutteellisen koulutuksen takia polttopuiden pilkkominen, jota ei pitänyt lukea tähän ryhmään, siitä huolimatta tuli siihen merkityksi suuremmissa määrin kuin B-tutkimuksessa. Toinen, yhtä todennäköinen syy voi olla siinä, että kun A-tutkimuksessa haastattelu kohdistui jokaiseen kalenterikuukauteen erikseen, haastattavallekin yleensä hämmästyttävän runsaat, mutta hajanaiset kotitarvepuun hakkuuseen ja ajoon käytetyt päivät ovat tulleet tarkemmin esille.

Todennäköistä on, että juuri viimeksi mainittu syy on samalla aiheuttanut sen, että ryhmässä »muu maataloustyö» ero B-tutkimuksen hyväksi on merkitsevä. A-tutkimuksessa yritettiin tarkoin juuri tästä työpanosryhmästä poistaa työttömyys-, loma-, hukka-, sairaus- ja opiskelupäivät sekä virkistysluonteiseenkin metsästyksen ja kalastukseen ynnä omaan koti- ja pienteollisuuteen käytetty työpanos. Kun tällaista »perkausta» ei suoritettu B-tutkimuksessa, on varsin luultavaa, että haastateltavat tietämättään tulivat siinä lukeneeksi tähän kaatoryhmään huomattavan osan perheensä mainitunlaisesta arkiajan käytöstä. Tämän takia taulukkoon 23/s. 117 A-tutkimuksen lukuihin on lisätty kyseisen henkilöryhmän metsästyksen, kalastukseen, poronhoitoon, omaan käsityöhön ja pienteollisuuteen käyttämä työpanos sekä sen avoimen työttömyyden aika ja hukkapäivät. Näin ero maatalojen koko oman työn panosta tarkoittavissa luvuissa pienenee niin huomattavasti, että sen merkitsevyys B-tutkimuksenkin varmuusvälin huomioon ottaen tulee varsin epävarmaksi.

Edellisen lisäksi on vielä huomattava, että A-tutkimuksessa työkyvyttömyiksi merkityt sekä 64 vuotta vanhemmat miehet osallistunevat muita miehiä suhteellisesti vähemmän raskaisiin metsä-, uudisrakennus- ja raiwaustöihin. Täten on luonnollista, että mainituissa ryhmissä työpanosluvut ovat molemmissa tutkimuksissa lähempänä toisiaan kuin ryhmässä »muu maatalous», johon keveimmät kotiaskareet sisältyvät.

Maaseudun työvoiman tutkimuksen ja maatalouslaskennan työvoimatiedustelun tulokset tärkeässä tilasuuruusluokassa 2—4.99 hehtaaria peltoa ovat siten varsin yhtäpitävät.

3352. Metsä- ja uittotyömiesten lukumäärä

Vuoden 1950 väestönlaskennassa oli jokaiselle ruokakunnalle jaetussa, ns. C-lomakkeessa jokaisen 14 vuotta täyttäneen ja sitä vanhemman henkilön annettava seuraava tieto: »Jos olette vuoden 1950 aikana olleet metsä- ja uittotöissä, ilmoittakaa näihin töihin yhteensä käyttämäne aika kuukausina. Omassa metsässä suoritettuja töitä ei ilmoiteta».

Tätä ilmoitusta käytettiin väestönlaskennan aineiston käsittelyssä luokiteltaessa 14 vuotta täyttäneet ja sitä vanhemmat ammattinharjoittajat elinkeinohaaran, ammattiaseman ja ammatin perusteella. Siten elinkeinohaaraan metsä- ja uittotyöt ja siinä ammattiasemaan toisen palkallisessa palveluksessa olevat työntekijät sekä ammattiryhmään metsä- ja uittotyöntekijät tuli lukea seuraavat henkilöt (Väestönlaskenta 1950, ohjeet):

a) Sellaiset henkilöt, jotka ilmoittivat pääammatikseen metsä- tai uittotyömies tai niitä vastaavan nimityksen.

b) Sellaiset henkilöt, jotka olivat ilmoittaneet käyttäneensä metsä- ja uittotöihin 4 kk. tai enemmän. Poikkeuksena olivat sellaisen 5 ha tai enemmän peltoa käsittävän maatilalan viljelijäperheen jäsenet, jotka kaikki ilmoittivat käyttäneensä metsä- ja uittotöihin 4 kk. tai enemmän. Heistä oli yksi merkittävä maanviljelijäksi (elinkeinohaara maatalous).

c) Edellisten lisäksi oli merkittävä metsä- ja uittotyöntekijöiden ryhmään sellaiset sekatyöntekijäksi pääammatinsa merkinneet, jotka olivat ilmoittaneet olleensa metsä- ja uittotöissä 2 tai 3 kk.

Väestönlaskennan (C-tutkimus) tulosten mukaan oli täten määritellen elinkeinohaaroissa metsätalous ja uitto ammattiasemaltaan toisen palkallisessa palveluksessa olevia työntekijöitä koko maan maalaiskunnissa ja Karhulan kauppalassa yhteensä 102 243 miestä (Tilastollisesta päätoimistosta saatu ennakkotieto).

A-tutkimuksessa oli metsä- ja uittotyömiehet mahdollisuus ryhmitellä palkatun metsä- ja uittotyöpanoksen perusteella 10-päiväisiin luokkiin. Metsä- ja uittotöitä vieraalle 71 päivää tai sitä enemmän tehneitä miehiä oli vastaavalla alueella A-tutkimuksen mukaan 118 817 miestä ja 81 päivää ja sitä enemmän ao. töitä tehneitä 99 337. Viimeksi mainittu luku on siis lähinnä C-tutkimuksen tulosta ja siihen parhaiten verrattavissa.

Olettaen, että kyseiset miehet keskimäärin tekevät näitä töitä n. 20 päivää kuukauden pituisena ajanjaksona, 4 kk. vastaa n. 80 työpäivää.

Taulukossa 24/s. 121 on lääneittäin verrattu näin määriteltyjen »ammattimaisten» metsä- ja uittotyömiesten määriä A- ja C-tutkimuksen (Tilastollisen päätoimiston antamia ennakkotietoja) perusteella. Turun ja Porin läänin, Ahvenanmaan, Vaasan läänin ja koko maaseudun luvuissa ei ole merkitsevää eroa. Muuten ovat erot erisuuruisia, jopa erisuuntaisia-kin. Näiden eroavuuksien merkitystä A-tutkimuksen tulosten luotettavuuden kannalta harkittaessa on huomattava, että C-tutkimuksen tulokset ovat olleet ainakin osaksi riippuvaiset kolmesta varsin sattumanvaraisesta seikasta.

Edellä esitetystä väestönlaskennan lomakkeen tiedustelun sanamuodosta ei käy ilmi, oliko merkittävä niiden kuukausien määrä, jolloin pääasiassa tehtiin mainittuja töitä, vai oliko mahdollisesti hajalliset työpäivät tai viikot laskettava yhteen ja summan perusteella arvioitava työpanoksen määrä kuukausina. A-tutkimuksessa kävi selville, että metsä- ja uittotyömiehillä metsä- ja uittotyöpanos yleisesti hajautuu vuoteen lukuisina lyhyinä jaksoina, joiden yhteistä pituutta tarkoin asiaa muistelematta on vaikea arvioida. Tiedusteltavat joutuivat säännön mukaan itse täyttämään väestönlaskennan C-lomakkeen. Heidän kyseiseen kohtaan merkitsemänsä vastaus C-tutkimuksessa riippui tällöin varsin olennaisesti siitä, miten he käsittivät kysymyksen sanamuodon ja siinä tapauksessa, että he pyrkivät ensin arvioimaan sanotun työpanoksensa määrän päivien summana, mitä jakajaa he arviossaan käyttivät. Tavanomaisin luku olisi ollut 25 työpäivää/kk. Täten väestönlaskennan lomakkeeseen tehty merkintä 4 kk. metsä- ja uittotöitä saattoi todellisuudessa vastata 60—100 päivän työpanosta.

Toisaalta C-tutkimuksen tulos tuli riippumaan siitä, kuinka moni henkilö, joka oli merkinnyt tehneensä kyseisiä töitä vähintään 2, mutta vähemmän kuin 4 kuukautta, merkitsi pääammatikseen sekatyömies tai jotain siihen verrattavaa. Tämä seikka puolestaan voi riippua sekä niiden henkilöiden lukumäärästä, joilla todella on useita tasa-arvoisia ammatteja, että väestön asennoitumisesta ammattinimityksiin yleensä. Molemmat seikat voivat olla sangen erilaiset eri läänien alueilla.

Kolmanneksi väestönlaskennan tulokset ovat riippuneet siitä, ketkä henkilöt ovat merkinneet pääammatikseen metsä- ja uittotyöt. On mahdollista, että ainakin jokin määrä miehiä, jotka ovat tehneet metsä- ja uittotöitä vähemmän kuin 81 työpäivää v. 1950, on siitä huolimatta merkinnyt pääammatikseen esim. metsätyömies. A-tutkimuksen koehenki-

Taulukko 24. »Ammattimaisten» metsä- ja uittotyömiesten lukumäärä lääneittäin. Päättökäytännön (A) ja väestönlaskennan (C) vertailu. Varmuusvälit on merkitty kaksinkertaisin keskivirhein.

Table 24. Number of paid »professional» forest and floating workers in the various counties. Comparison between Main Survey (A) and Population Census of 1950 (C). Confidence intervals with doubled standard errors.

Lääni — County	A Miehiä Males	C Miehiä Males	$\frac{A}{C}$ %
Uudenmaan — Uusimaa	3 080	2 479	124 ± 19
Turun ja Porin — Turku and Pori	6 836	6 384	107 ± 11
Ahvenanmaa — Åland	53	53	100 ± 116
Hämeen — Häme	11 135	9 372	119 ± 10
Kymen — Kymi	5 648	4 695	120 ± 13
Mikkelin — Mikkelä	13 063	11 648	112 ± 8
Kuopion — Kuopio	23 480	26 581	88 ± 5
Vaasan — Vaasa	14 672	14 359	102 ± 7
Oulun — Oulu	13 147	16 499	80 ± 6
Lapin — Lapland	8 223	10 173	81 ± 7
Koko maaseutu — All rural districts	99 337	102 243	97 ± 3

löissä oli metsätyömieheksi itseään nimittäneitä, jotka eivät ensinkään olleet osallistuneet ansiotöihin metsässä v. 1950.

Vielä on huomattava, että A-tutkimuksen lukuihin sisältyvät vain 15—64-vuotiaat, mutta C-tutkimuksen lukuihin 14 vuotta täyttäneet aommatin harjoittajat.

Jos taulukon 24 lukuja tarkastellaan edellä esitetyn valossa, on ajateltavissa, että $\frac{A}{C}$ on pienempi kuin 100 % niissä lääneissä, joissa on ollut runsaasti sekatyömieheksi itseään nimittäneitä metsätyöntekijöitä. Kuopion, Oulun ja Lapin lääneissä heitä on voinut olla runsaammin kuin maan keskiosissa ja rannikkoalueilla. Uudenmaan, Hämeen, Kymen ja Mikkelin lääneissä sanottu suhdeluku merkitsevästi ylittää 100 %. Tämä on voinut johtua siitä, että näillä alueilla C-tutkimuksessa työpanoksen 4 kk. metsä- ja uittotöitä on yleisesti katsottu merkitsevästi suurempaa työpanosta kuin 80 työpäivää. On mahdollista, että rannikolla ja Hämeessä näin on ollut asianlaita. Toisaalta on myös todennäköistä, että kun puheena olevilla alueilla A-tutkimuksen mukaan viljelijöiden keskimääräinen lukumäärä

5—10 peltohehtaarin maatilaa kohti on pienempi kuin Itä- ja Pohjois-Suomessa, 4 kk. ja enemmän metsä- ja uittotöitä tehneitä miehiä väestönlaskennan käsittelyssä on merkitty maanviljelijöiksi suhteellisesti runsaimmin maan rannikkoseuduilla ja Hämeen läänissä. Mikkelin läänin suhdeluku $\frac{A}{C} = 112\%$ ei kuitenkaan voi täten selittyä, koska ko. lääni näissä suhteissa muistuttaa lähinnä Kuopion läänää.

A- ja C-lukujen väliset erot voivat, etenkin Uudenmaan, Hämeen, Kymen ja Mikkelin lääneissä johtua myös A-tutkimuksen näytteen harhaisuudesta. Vakuuttavaa todistetta tästä seikasta esitetty vertailu ei anna. Lähinnä se pystyy vain osoittamaan, *ettei A- ja C-tutkimusten edellä selostetuissa tuloksissa ole selvästi niin suuria eroavuuksia, etteivät ne voi johtua huomattavasti toisistaan poikkeavista mittaus- ja tulosten käsittelymenetelmistä.*

3353. Metsä- ja uittotyöpanoksen suuruus ja kausivaihtelu

Kululaitosten ja yleisten töiden ministeriö kerää kuukausittain metsätyönvälitystoimistojensa kautta tietoja metsä- ja uittotyöntekijöiden ja näistä erikseen hakkuumiesten lukumäärästä (D-tutkimus). Tiedot tarkoitavat tilannetta kunkin kalenterikuukauden päättyessä, ja ne yritetään saada kaikilta metsä- ja uittotyönantajilta. Saatu aineisto ei ainakaan sanottavasti peitä yksityismetsänomistajien hankintahakkuuta ja -ajoa eikä kotitarvepuun hakkuuta ja ajoa. Luultavasti myös osa muista järjestäytymättömistä työnantajista jää aineiston ulkopuolelle. Vuoden 1950 alkutiedot eivät myöskään peittäneet maan keski- ja rannikkoalueita, vaan metsätyönvälitystoimistojen toiminta-alueilta saadut luvut suurennettiin koko maata koskeviksi 1940-luvun alkupuoliskon »täydellisten» tilastojen perusteella arvioidulla kertoimella 2. Tulokset julkaistaan Valtiovarainministeriön kansantalousosaston suhdannesarjoissa.

Mainituista työvoimatiedoista voidaan työntekijöiden ansiotyöpanos metsä- ja uittotöissä arvioida siten, että esim. joulukuun lopussa 1949 ja tammikuun lopussa 1950 saatujen työntekijämäärien keskiarvo kerrotaan luvulla 19, joka suunnilleen vastaa yhden metsätyömiehen keskimääräistä metsätyöpanosta kuukaudessa (PÖNTYNEN 1936, s. 48). Saatu luku on silloin metsätyöpanoksen summa tammikuussa 1950.

D-tutkimuksen metsä- ja uittotyömiesten kuukausittaisista kokonaisu-määrästä täten lasketut työpanosluvut on merkitty taulukkoon 25/s. 123,

Taulukko 25. Ansiotyöpanos metsätaloudessa ja uittossa, kotitarvepuun hakkuuta ja ajoa lukuun ottamatta, kuukausittain, koko maaseutu. Pää-tutkimuksen (A) sekä kululaitosten ja yleisten töiden ministeriön tilaston (D) vertailu. Varmuusvälit on merkitty kaksinkertaisin keskivirhein.

Table 25. Input of paid labour to forestry and floating, excl. felling and hauling of wood for domestic use. Monthly figures for all rural districts. Comparison between Main Survey (A) and statistics of Ministry of Communications and Public Works (D). Confidence intervals with doubled standard errors.

Kuukausi Month	A 1 000 miespäivää man-days in thousands	D 1 000 miespäivää man-days in thousands	¹ A D %	² A %	D %
I	2 963	2 062	144 ± 3	16	14
II	3 124	2 613	120 ± 2	17	18
III	2 573	1 805	143 ± 3	13	12
IV	1 523	1 093	139 ± 4	8	7
V	1 215	950	128 ± 4	6	6
VI	1 020	903	113 ± 4	5	6
VII	611	646	95 ± 5	3	4
VIII	634	456	139 ± 7	3	3
IX	684	551	124 ± 6	4	4
X	1 080	789	137 ± 5	6	5
XI	1 752	1 235	142 ± 4	9	8
XII	1 998	1 853	108 ± 3	10	13
Koko vuosi — Total year 1950	19 177	14 956	128 ± 2	100	100

ja niitä on verrattu A-tutkimuksen vastaaviin lukuihin. Vertailukelpoisuutta vähentää D-tutkimuksen puutteellisen peittävyuden lisäksi se, että A-tutkimuksen lukuihin taulukossa ei sisälly metsänomistajaperheiden työpanos omissa metsissä silloinkaan, kun se on tehty puiden pystyyn ostajien palkkaamana. Sen sijaan se sisältynee D-tutkimuksen lukuihin.

¹ Koska vajaamuotoisen D-tutkimuksen varmuusrajoja ei tunneta, ei niitä ole voitu ottaa huomioon. Todellisuudessa ovat siten varmuusvälit sarakkeeseen merkityjä huomattavasti laajemmat — *As the confidence limits of the partial D-survey are not known they cannot be taken into account. Thus the true confidence intervals are considerably larger than those marked in the column.*

² Varmuusvälit ovat pienemmät kuin ± 0.5 sadannesyksikköä — *Confidence limits are less than ± 0.5 percentage points.*

Taulukko 26. Ansiotyöpanos markkinapuun hakkuussa, kuukausittain, koko maaseutu. Päättökäytön (A) sekä kulkulaitosten ja yleisten töiden ministeriön tilaston (D) vertailu. Varmuusvälit on merkitty kaksinkertaisin keskivirhein.

Table 26. Input of paid labour to felling of wood for sale. Monthly figures for all rural districts. Comparison between Main Survey (A) and statistics of Ministry of Communications and Public Works (D). Confidence intervals with doubled standard errors.

Kuukausi Month	A 1 000 miespäivää man-days in thousands	D 1 000 miespäivää man-days in thousands	¹ A D %	² A %	D %
I	1 682	1 187	142 ± 4	15	15
II	1 723	1 378	125 ± 4	16	17
III	1 427	903	158 ± 5	14	11
IV	831	380	219 ± 9	8	5
V	418	190	220 ± 12	4	2
VI	376	257	146 ± 9	4	3
VII	218	276	79 ± 6	2	3
VIII	289	266	109 ± 8	3	3
IX	393	390	101 ± 6	4	5
X	716	618	116 ± 5	7	8
XI	1 195	950	126 ± 4	11	12
XII	1 293	1 331	97 ± 3	12	16
Koko vuosi — Total year 1950	10 561	8 126	130 ± 3	100	100

Lisäksi A-tutkimuksessa todettiin metsätyömiesten keskimääräisen metsätyöpanoksen eri kuukausina vaihtelevan, mitä lukuja laskettaessa ei kuitenkaan voitu ottaa huomioon. Myöskään metsä- ja uittotöiden kausivaihtelu Itä- ja Pohjois-Suomessa, josta D-tutkimuksen aineisto on kerätty, ei ole samanlainen kuin maan muissa osissa. Kausivaihtelun erilaisuus ei siten ole päässyt vaikuttamaan D-tutkimuksen perusteella laskettuihin lukuihin. Lopuksi todettakoon, että D-tutkimuksen lukujen laskumenetelmä perustuu olettamukseen, että metsä- ja uittotyömiesten lukumäärä kuukauden alusta sen päättymiseen kehittyy suoraviivaisesti, mikä ei suinkaan aina pidä paikkaansa (vrt. kuvaa 2/s. 50).

Kun mainituin varauksin tarkastellaan taulukon 25 lukuja $\frac{A}{D}$, ne

^{1, 2} Ks. vastaavia alaviittoja taulukossa 25/s. 123 — See foot-notes in Table 25/p. 123.

heinäkuuta lukuun ottamatta osoittavat melkoista yhdenmukaisuutta vuoden eri kuukausina. Täten myös A- ja D-tutkimusten kausivaihtelu on verrattain samanlainen.

Taulukkoon 26/s. 124 on vastaavalla tavalla merkitty markkinapuun hakkuutyöpanos kuukausittain A- ja D-tutkimusten perusteella. Tällä kertaa tulosten yhdenmukaisuus on huomattavasti pienempi kuin edellä. Varsinkin huhti- ja toukokuun sekä heinäkuun suhdeluvut poikkeavat hyvin paljon koko vuoden keskiarvosta. D-tutkimuksen mukaan työpanos toukokuussa olisi ollut vuoden pienin, vieläpä paljon pienempi kuin heinäkuussa, jolloin hakkuutyöt heinänteon ja puutavaran suuren pilaantumisvaaran takia yleensä ovat hyvin vähäiset. Tällainen tulos lienee epätodennäköisempi kuin A-tutkimuksen osoittama. Myöskään D-tutkimuksen osoittama jyrkkä töiden väheneminen maaliskuusta huhtikuuhun ei tunnu todennäköiseltä. Täten tulosten eroavuus tuskin voi osoittaa A-tutkimuksen tuloksia harhaisiksi.

3354. Työttömyysajan määrä ja kausivaihtelu

Suomen virallinen työttömyystilasto kootaan siten, että kunnat ilmoittavat työvoimapiirien välityksellä kunkin kalenteriviikon lopussa kulkulaitosten ja yleisten töiden ministeriölle työttömyyskortistossa sanottuna päivänä olleiden työttömyystöihin sijoitettujen ja niihin sijoittamattomien henkilöiden määrän, miehet ja naiset erikseen. Tilasto julkaistaan Sosiaalisessa aikakauskirjassa.

A-tutkimuksen tuloksiin verrattavien lukujen saamiseksi arvioitiin kulkulaitosten ja yleisten töiden ministeriön luovuttamasta alkuaineistosta maalauskuntien työttömien miesten työttömyystyöpanos sekä avoimen työttömyyden aika miespäivinä kuukausittain v. 1950. Viikon päättyessä kortistoihin merkittyjen miesten lukumäärä kerrottiin sanotun viikon arkipäivien lukumäärällä ja tulot laskettiin yhteen kalenterikuukausittain. Milloin kalenteriviikko jakautui kahden kuukauden osalle, jaettiin sen viikon tulo molempien kuukausien osalle niihin kuuluvien arkipäivien osoittamissa suhteissa. Tässäkin arvioimismenetelmässä syntyy virheitä toisaalta sen tähden, että kaikki työttömyystöissä olevat miehet eivät ole työssä eivätkä myöskään töihin sijoittamattomat miehet työttöminä jokaisena arkipäivänä, ja toisaalta sen tähden, että miesten lukumäärä lasketaan viikon päättyessä eikä sen keskikohtalla. Kun arviointiajanjakso on hyvin lyhyt, jälkimmäiset virheet ovat kuitenkin pieniä. Tämän ohessa

Taulukko 27. Maaseudun miesten työttömyystyöpanos työvoimapiireittäin. Päättökimukseen (A) ja kuntien työttömyyskortistojen (E) vertailu. Varmuusvälit on merkitty kaksinkertaisin keskivirhein.

Table 27. Labour input to relief work in the various employment districts. Rural manpower. Comparison between Main Survey (A) and communal unemployment registers (E). Confidence intervals with doubled standard errors.

Työvoimapiiri Employment district	A 1 000 miespäivää man-days in thousands	E 1000 miespäivää man-days in thousands	$\frac{A}{E}$ %
Helsingin — Helsinki	15	37	41 ± 27
Turun — Turku	62	89	70 ± 22
Tampereen — Tampere	107	175	61 ± 15
Jyväskylän — Jyväskylä	70	160	44 ± 13
Kouvolan — Kouvola	121	253	43 ± 11
Kuopion — Kuopio	213	431	49 ± 8
Joensuun — Joensuu	236	399	59 ± 9
Vaasan — Vaasa	108	177	61 ± 15
Kajaanin — Kajaani	124	189	66 ± 15
Oulun — Oulu	726	1 274	57 ± 6
Rovaniemen — Rovaniemi	233	495	47 ± 8
Koko maaseutu — All rural districts	2 015	3 679	55 ± 3

tulee lopputuloksiin virheitä myös sen tähden, että kunnat joskus jättävät ilmoitusvelvollisuutensa täyttämättä.

Taulukkoon 27 yllä on työvoimapiireittäin merkitty mainitulla tavalla arvioitu maalaiskuntien miesten työpanos työttömyystöissä työttömyyskortistojen mukaan (E) ja verrattu sitä A-tutkimuksen vastaaviin tietoihin. Vertailun tulos osoittaa, että A-tutkimuksen tulokset ovat kauttaaltaan niin paljon pienemmät kuin E-tutkimuksesta saadut luvut, että *eron selvästi täytyy johtua A-tutkimuksen harhasta*. Harhan syy on ilmeisesti siinä, että työttömyystyöpanosta tiedusteltiin maaseudun työvoiman tutkimuksen T-lomakkeessa liian vaikeatajuisella tavalla ja lisäksi haastattelun lopussa, jolloin sen kohteen täytyi jo olla varsin väsynyt lukuisiin tiedusteluihin. Juuri työttömyystyöpanostiedoissa havaittiin T-lomakkeissa runsaimmin merkintävirheitä, jotka osoittivat haastattelijoiden usein ymmärtäneen kysymyksen väärin. Esitutkimuksissa voitiin myös havaita, että haastateltavien oli monesti varsin vaikea muistaa, milloin

Taulukko 28. Maaseudun miesten työttömyystyöpanos kuukausittain, koko maaseutu. Päättökimukseen (A) ja kuntien työttömyyskortistojen (E) vertailu. Varmuusvälit on merkitty kaksinkertaisin keskivirhein.

Table 28. Labour input to relief work, all rural districts, monthly figures. Comparison between Main Survey (A) and communal unemployment registers (E). Confidence intervals with doubled standard errors.

Kuukausi Month	A 1 000 miespäivää man-days in thousands	E 1 000 miespäivää man-days in thousands	$\frac{A}{E}$ %
I	363 ± 23	740	49 ± 3
II	431 ± 24	792	54 ± 3
III	505 ± 27	978	52 ± 3
IV	355 ± 21	734	48 ± 3
V	167 ± 15	325	51 ± 5
VI	67 ± 10	26	258 ± 37
VII	19 ± 5	—	.
VIII	16 ± 5	—	.
IX	14 ± 5	0	.
X	20 ± 5	13	154 ± 42
XI	28 ± 6	25	112 ± 26
XII	30 ± 6	46	65 ± 14
Koko vuosi — Total year 1950	2 015 ± 112	3 679	55 ± 3

jokin työmaa oli ollut työttömyystyömaa ja milloin se taas oli avoin työmaa. Lisäksi pienten ajankäyttöryhmien kato vaikutti työttömyystyöpanosta pienentävästi.

Taulukossa 28 yllä sama vertailu on vielä tehty kuukausittain. Vertailun tulos osoittaa, että niinä kuukausina, jolloin työttömyystyömaita on ollut suhteellisen runsaasti käynnissä, A-tutkimuksen luvut ovat keskimäärin noin puolet E-tutkimuksen työpanoslukuista. Sen sijaan kesä-marraskuussa, jolloin työttömyystöitä todellisuudessa on teetetty vähän tai ei ollenkaan, A-tutkimuksessa on mittausvirheistä johtuen saatu suhteellisesti paljon suurempia, joskin absoluuttisesti vähämerkityksisiä työpanoslukuja.

Vertailu A- ja E-tutkimusten avoimen työttömyysajan määrien välillä osoitti, että esim. koko vuoden avoimen työttömyyden aika koko valtakunnassa oli A-tutkimuksessa noin 4 kertaa niin suuri kuin työttömyyskor-

tistosta saatu vastaava luku. Koska tämä tulosten kannalta hyvin mielenkiintoinen seikka ei kuitenkaan sanottavasti voi valaista A-tutkimuksen luotettavuuskysymystä, jätetään se tässä käsittelemättä.

34. Päättökäytön menetelmän yleisarvostelu

Vuoden 1950 maaseudun työvoiman tutkimusta varten poimittiin systemaattinen näyte väestönlaskennan yhteydessä kaikista maalaiskunnissa 31. 12. 1950 asuneista 15—64-vuotiaista miespuolisista henkilöistä eräitä laitoshoidossa olevia ryhmiä lukuun ottamatta. Tutkimuksen tulokset ja aineiston käsittelyssä saatu kokemus osoittavat, että näytteen tulisi peittää kaikki maalaiskunnissa asuvat yli 13-vuotiaat miespuoliset henkilöt, suurentamisen helpottamiseksi mieluummin myös laitoshoidokit, jotta asetettu päämäärä täysin saavutettaisiin.

Näytteen poiminnassa käytettiin satunnaisnumeroita vastaavina näyttekäytöiden tunnuksina syntymäpäiviä 1 ja 15 siten, että syntymäpäivä perustui etupäässä koehenkilön omaan ilmoitukseen. Poimintamenetelmä on periaatteessa käyttökelpoinen, mutta syntymäpäiviä valittaessa on otettava huomioon varsinkin syntymäpäivän 1 ja luultavasti kaikkien 5:llä jaollisten syntymäpäivien kasautumistaipumus, joka vaikeuttaa näytteen tunnuslukujen suurentamista ja lisäksi voi aiheuttaa vaikeasti hallittavia harhoja. Mikäli käytetään syntymäpäiviä, joihin kasautumistaipumus ei vaikuta, ja vastausten kadon osuus on määritettävissä, on näin poimitun näytteen suurentaminen mahdollista perusjoukon suuruutta tuntemattakin käyttämällä syntymäpäivien lukumääristä laskettua teoreettista suurennustekijää.

Mittausmenetelmänä käytettiin kokonaisen vuoden arkiajan käytön taannehtivasti haastattelemista siten, että mittayksikköinä olivat arkipäivät ja niiden jaottelu ajankäyttöryhmiin suoritettiin kalenterikuukausittain. Mittausmenetelmän luotettavuudesta suoritettujen tutkimukset osoittivat, että vastauksissa annetut luvut yksityistapauksissa voivat huomattavasti poiketa todellisista ajankäytön määristä. Yleensä nämä poikkeamat ovat satunnaisia, ja niiden vaikutus arviolukujen varmuuteen on siten keskivirhelaskelmissa todettavissa. Kuitenkin esiintyy pienten ajankäyttöryhmien systemaattista katoa. Tämä vastausharha pienentää tuloksissa muun arkiajan käytön kuin työpanoksen määrää, etenkin työttömyysaikaa sekä mahdollisesti pienimpiä työpanosryhmiä. Varsinkin oman työn panoksen ryhmiä se absoluuttisesti suurentaa, mutta suhteelli-

nen harha on tarkistuksista päätellen niissä vähäpätöinen. Koska edellä mainitusta huolimatta avoimen työttömyysajan määrä päätökäytön mukaan oli noin 4 kertaa niin suuri kuin työttömyyskortistosta laskettu vastaava luku, osoittanee tämä, että vastausharhan vaikutus ei mainittavasti voi vähentää tulosten käyttökelpoisuutta.

Vastausten kadon aiheuttamaa näytteen harhaisuutta tutkittiin vain ikäjakautumassa ja ammattijakautumassa. Näytteen ammattijakautuma on ainakin Uudenmaan läänissä, jossa vastausten kadon osuus oli poikkeuksellisen suuri, maatalouden ja metsätalouden ammattiryhmiä suosiva. Uudenmaan läänin näytteestä laskettuja tunnuslukuja on sen tähden syytä käyttää varovasti. Lapin läänin näytteessä ei havaittu harhaisuutta. Muiden läänien näytteiden harhaisuutta ei tutkittu, mutta koska niissä vastausten kato oli suhteellisesti pienempi kuin mainituissa lääneissä, voi olettaa, että niiden kadosta johtuvat harhat eivät ylitä kaksinkertaisen keskivirheen rajoja. Vastausten suuri kato johtui ennen kaikkea väestönlaskennalle asetetusta liian suuresta rasiuksesta. Tässä mielessä tietojen kerääminen työvoimatutkimusta varten on edullisempaa suorittaa erillisesti.

Tulosten mahdollisia harhoja pyrittiin tutkimaan myös vertaamalla laskettuja arvioita toisten tutkimusten vastaaviin tuloksiin. Työttömyys-työpanoksessa todettiin harha, joka tuloksissa korjattiin. Muissa vertailuissa todetut epävarmuudet johtunevat pikemminkin vertailukohteen kuin päätökäytön harhaisuudesta.

Analysoimalla satunnaisvaihtelun vaikutusta tulosten varmuuteen todettiin, että keskivirheet ovat riittävällä tarkkuudella laskettavissa arvioluvuista.

Tutkimuksessa käytetty haastattelulomake sekä aineiston käsittelyä varten kehitetty reikäkorttimenetelmä täyttivät tyydyttävästi niille asetetut vaatimukset. Niiden yksityiskohdissa on kuitenkin saatujen kokemusten perusteella tehtävissä korjauksia.

Näytteestä laskettujen arviolukujen varianssikaavoista voidaan, ratkaisemalla ne näytteen suuruuden n suhteen, arvioida määrättyä varmuusvaatimusta varten vaadittavan yksiaasteisen, osittamattoman otoksen suuruus. Ratkaisemalla n kaavasta 12/s. 114 saadaan

$$n = \frac{1}{v_{x_j}^2} \left[\frac{(1 - \varphi)}{\varphi} + \frac{c^2 (1 - \bar{p}_j)}{\varphi \bar{p}_j} \right] \quad (15)$$

Jos tutkimuksen kohteena ovat kaikki työkykyiset miehet ja näiden osuus näytteestä on 90 % (kun näytteeseen otetaan kaikki 14 vuotta täytäneet ja sitä vanhemmat miehet, on työkyvyttömiä määrä suhteellisen suuri), on $\varphi = 0.9$. Ajankäyttörhymälle, jonka \bar{p}_j -estimaatti on 0.1, saadaan c -kertoimella 0.75 ja varmuusvaatimuksella $v_{x'_j} = 0.05$ vaadittavan näytteen suuruudeksi $n = 2\,284$. Kun tämä näyte ajankäyttörhymissä, joiden $\bar{p}_j > 0.1$, antaa suhteellisesti pienemmän keskivirheen ja ajankäyttörhymälle, jonka \bar{p}_j -arvio = 0.01, saadaan vielä $v_{x'_j} = 0.16$, voidaan näytettä pitää käytännössä hyvin riittävänä.

Esitettyä suuruusluokkaa olevan otoksen riittävyys edellyttää kuitenkin, että todellinen suurennustekijä on laskettavissa perusjoukon ja näytteen suuruuden suhteena. Jos perusjoukon suuruutta ei tunneta, vaan tyydytään teoreettiseen suurennustekijään (poiminnassa käytetään syntymäpäiviä), on lisäksi otettava huomioon sen satunnaispoikkeamat todellisesta suurennustekijästä. Teoreettisella suurennustekijällä suurennettua näytteen n' (näytteestä lasketun perusjoukon suuruuden arvion) suhteellinen keskivirhe saadaan silloin kaavasta

$$v_{n'}^2 = \frac{1-f}{nf},$$

jossa f on teoreettinen otantaosuus $\frac{n}{n'}$. Ratkaisemalla n saadaan

$$n = \frac{\sqrt{4v_{n'}^2 n' + 1} - 1}{2v_{n'}^2}$$

Kun Suomen maaseudulla n' on suuruusluokkaa $8 \cdot 10^5$, tulee varmuusvaatimuksella $v_{n'} = 0.05$ näytteen suuruudeksi $n \sim 18\,000$ koehenkilöä, joten ensiksi mainittu varmuusvaatimus likimäärin tulisi täytetyksi koko maaseudun arvioluvuissa vasta $n. 20\,000$ koehenkilön näytteellä. Laskuesimerkki osoittaa, että *otantaosuuden ollessa hyvin pieni teoreettisen suurennustekijän käyttäminen ei ole taloudellista pyrittäessä absoluuttisiin arviolukuihin*, jos perusjoukon todellinen suuruus edes kohtalaisin kustannuksin saadaan selville. Väestönlaskennan ja henkikirjoituksen perusteella se voitaneen useimmissa tapauksissa saada verrattain helposti lasketuksi tai ainakin suurella varmuudella arvioiduksi. Jos tyydytään pelkästään suhdelukuihin \bar{p}_j , ei perusjoukon suuruuden tunteminen ole tarpeen. Suhdelukujen \bar{p}_j arviointia varten vaadittavan otoksen suuruus on

$$n = c^2 \bar{q}_j / v_{\bar{p}_j}^2 \varphi \bar{p}_j \quad (16)$$

joten samoin varmuusvaatimuksin kuin edellä, saadaan $n = 2\,240$.

Laskuesimerkkien osoittama näyte vaaditaan, kun halutaan arvioida koko valtakunnan maaseudun miestyövoiman käyttö kokonaiselta vuodelta. Kuukausien arvioluvuilla x'_j ja \bar{p}_j tulee varmuus samansuuruisella näytteellä hiukan pienempi, $v_{(\bar{p}_j = 0.1)} = \pm 0.06$ ja estimaateille μ' ja ψ vielä hiukan pienempi, esim. $v_{(\psi = 0.1)} = \pm 0.07$, jos φ edelleen on 0.9. Läänin, metsänhoitolautakunnan toiminta-alueen ja työvoimapiirin tunnuslukujen arvioimiseksi samalla tarkkuudella tarvitaan likimäärin samansuuruisen näyte. Kun vuoden 1950 maaseudun työvoiman tutkimuksessa laskettiin edellä mainitut tunnusluvut sangen pienillekin otoksen alaryhmille ja lisäksi muutamien harvojen kuntien muodostamille pienalueille, ei sen näytettä voida pitää tähän tarkoitukseen liian suurena. Jos vastaisuudessa samanlaatuisia tutkimuksia suoritetaan suhteellisen lyhyin väliajoin kehityssuunnan selvittämiseksi, voitaneen useimmiten tyytyä koko valtakuntaa koskeviin tietoihin. *Mikäli osittamattoman, yksiasteisen otoksen käyttö on mahdollista, on n. 2\,500—3\,000 koehenkilön näyte silloin riittävä.* Jos perusluettelo (henkikirja) voidaan poimintaa suoritettaessa jakaa kahteen ositteeseen siten, että henkisen alan ja teollisuuden työlliset tulevat toiseen ja kaikki muut perusyksiköt toiseen ositteeseen, lisääntynee tulosten varmuus jonkin verran. Etenkin keskivirhelaskelmien luotettavuus suurenee (COCHRAN 1953, ss. 27—30).

35. Päättökustannukset

Yhdistyneitten Kansakuntien tilastotoimiston suositusten mukaan (The preparation ... 1950, s. 11) tulee otantatutkimuksen selostukseen liittää tiedot kustannuksista. Vaikka tätä ohjetta harvoin on noudatettu, sen täyttäminen on helposti perusteltavissa. Nykyaikaisten otantatutkimusten poimintamenetelmää suunniteltaessa on kustannustekijällä ratkaiseva merkitys sekä poiminta-asteiden lukumäärän ja laadun että näytteen suuruuden määrittämisessä (HANSEN ym. 1953, ss. 270—283). Näin ollen vastaisuudessa menetelmältään samantapaista tutkimusta suunniteltaessa aikaisemman tutkimuksen kustannuksista saadut tiedot voivat olla avuksi. Sitä paitsi seuraavassa esitettävillä aineiston käsittelyn yksikkökustannusluvuilla voi olla merkitystä myös muiden tutkimusten käsittelykustannuksia arvioitaessa.

Maaseudun työvoiman tutkimuksen suorittamista varten Eduskunta myönsi lisäyksenä vuoden 1951 tulo- ja menoarvioon (15 Pl. 55. mom.) 8 441 800 markan siirtomäärärahan, josta 31. 12. 1953 mennessä, jolloin aineiston käsittelytyöt saatettiin katsoa loppuun suoritetuiksi, oli käytetty 8 165 364 markkaa. Menojen jakautuminen eri käyttötarkoituksiin selviää seuraavasta asetelmasta.

	Yhteensä mk	Mk/yksikkö	
Vakinaisen henkilökunnan palkat	1 779 675: —		
Aineiston keräyspalkkiot	3 449 525: —	75: —/lomake	} 86: 48/lomake
Perusaineiston esitarkistus, palkat	242 180: —	5: 06/lomake	
Perusaineiston laskenta ja kodifointi, palkat	307 545: —	6: 42/lomake	} 2: 93/ kortti
Henkilökorttien lävistys, palkat	105 490: —	2: 25/kortti	
Henkilökorttien lävistys, konekustannukset	31 843: —	0: 68/kortti	} 1: 57/kortti
Aikakorttien käsin merkinnät, palkat	371 403: —	1: 24/kortti	
Aikakorttien lävistys, konekustannukset	98 003: —	0: 33/kortti	
Lajittelu-, yhteenlasku-, toimisto-, kertomis- ja taulukointityöt. Palkat ja koneiden vuokrat	218 169: —		
Henkikirjojen laskentatyöt, palkat	171 531: —		
Lomakkeet, reikäkortit ym. tarvikkeet	866 877: —		
Kuljetus-, postitus- ja järjestelytyöt yms.	158 358: —		
Työnantajan KEL- ja lapsilisämaksut	344 765: —		
	8 165 364: —		

Koska menot joitakin vähäisiä summia lukuun ottamatta on suoritettu vuosina 1951—1953, jolloin palkka- ja hintataso pysyi jokseenkin muuttumattomana (kyseisten töiden kannalta), ovat kustannuserät sikäli vertailukelpoiset. On huomattava, että tutkimuksessa toiminut vakinainen henkilökunta sai pääosan palkastaan Helsingin yliopistolta, jonka tehtäviä se myöskin oli velvollinen hoitamaan tutkimustyön ohella. Tästä johtuen tutkimuksen johtokustannukset eivät sisälly kokonaisuudessaan asetelmassa mainittuihin summiin, eikä niiden todellista suuruutta ole helppo myöskään arvioida. Laitoksen vakinainen henkilökunta suoritti käsittelytyön valvonnan, esitarkistuksen tarkistuksen, jälkitarkistustyöt, huomattavan osan lajittelutyöstä, aikakorttien koneellisen lävistyksen, kertomistyöt sekä kaikki koneellisesti taulukoitujen aineistojen jälkikäsitteilyt. Sitä paitsi sen tehtäviin sisältyivät kaikki toimistotyöt tutkimuksessa, kuten palkanmaksu- ja puhtaaksikirjoitus. Tämä on otettava huomioon

verrattaessa henkilökorttien lävistyskustannuksia (2: 93/reikäkortti) aikakorttien lävistyskustannuksiin (1: 57/reikäkortti). Jälkimmäinen työ tuli joka tapauksessa huomattavasti halvemmaksi korttia kohti kuin edellinen, vaikka aikakorttien mark sensing-merkintä ja niiden koneellinen lävistys oli huonojen reikäkorttien ja koneen heikon kunnon takia pääosaltaan suoritettava kahteen kertaan.

Perusaineiston keräyspalkkioita useimmat laskentatoimikunnat pitivät riittävinä. Jos aineiston keräys olisi tapahtunut erillisenä eikä väestönlaskennan yhteydessä, keräyskustannukset olisivat tulleet näytekysikköä kohti huomattavasti suuremmiksi.

Aineiston keräyksen vaatimasta työajasta maaseudun työvoiman tutkimuksessa ei kerätty tietoja. Sen sijaan Ranuan ajankäyttötutkimuksessa kesällä 1949 merkittiin haastatteluun kulunut aika lomakkeeseen. Yhteensä 560:sta tällä merkinnällä varustetusta haastattelusta 47 % oli vaatinut 20—30 minuuttia. Keskimääräinen aika oli 24 minuuttia. Alle 10 tai yli 40 minuuttia kestäviä haastatteluja oli molempia alle 10 % koko määrästä. Haastatteluajan pituutta arvioitaessa on huomattava, että Ranualla kukin haastattelijä joutui suorittamaan haastatteluja varsin runsaasti (toiset jopa satoja), joten niihin kuluva aika tottumuksen ansiosta tuli lyhimmäksi mahdolliseksi. Toisaalta on muistettava, että Ranualla haastateltavat olivat miltei poikkeuksetta runsaasti aikaa vieviin tapauksiin luettavia.

Aineiston käsittelyssä käytetyn tilapäisen henkilökunnan palkkauksessa (perusaineiston esitarkistus, laskenta ja kodifointi, aikakorttien käsin merkintä sekä henkikirjojen laskenta) pyrittiin työ suorittamaan mahdollisimman suuressa määrin yksikköpalkalla. Kokemukset tästä urakkatyöstä olivat myönteisiä. Kun työn laadun tarkkailu järjestettiin tehokkaaksi ja virheellisen työn tekijät parin varoituksen jälkeen viipymättä erotettiin, saatiin tarjolla olleesta runsaasta työvoimasta valituiksi nopeat ja samalla huolelliset työntekijät. Yksikköpalkkauksen ansiosta he saivat myös kykyjään vastaavan palkan työstään. Samalla yksikkökustannukset alussa kokeiltuun tuntipalkkajärjestelmään verrattuna laskivat huomattavasti. Varsinkin tilapäistä (opiskelija-)työvoimaa käytettäessä, jolloin työajan tarkkailua on vaikea järjestää ja ankara laadun tarkkailu joka tapauksessa on tarpeellista, yksikköpalkkausta on pidettävä tämänlaatuisissa töissä suositeltavana.

Edellä selostettujen kustannusten suuruutta arvioitaessa on huomattava, että niihin sisältyvät kaikkien saman aineiston perusteella julkaistujen tai julkaistaviksi suunniteltujen tutkimusten aineiston keräys- ja käsittelykustannukset.

4. Työpanoksen, työttömyysajan ja työvoiman rakenteen tutkimusmenetelmät saatujen kokemusten valossa

Vuoden 1950 maaseudun työvoiman tutkimuksen tarkoituksena oli saada riittävän luotettava kuva kolmesta, luonteeltaan toisistaan poikkeavasta tutkimuskohteesta:

1) Maaseudun miespuolisen väestön työpanoksesta havaintovuoden aikana, sen jakautumisesta yhteiskuntaryhmittäin, elinkeinoittain ja alueittain sekä työpanoksen rakenteen kausivaihtelusta.

2) Maaseudun miespuolisen väestön työttömyysajasta, myös luppoajasta eli piilevästä työttömyyden ajasta ja sen kausivaihtelusta.

3) Eräissä maaseudun elinkeinohaaroissa, etenkin maa- ja metsätaloudessa havaintovuoden aikana lyhyemmän tai pitemmän ajan toimineiden miespuolisten työllisten lukumääristä, näiden työvoimaryhmien koostumuksesta ja pää- tai sivuammattin merkityksestä työllisten arkiajan käytössä.

Jonkin tuotannonalan, esim. metsätalouden *työpanos* on mitattavissa siten, että tietojen alkulähteinä ovat joko yritykset (työnantajat) tai työn suorittajat (työlliset). Edellistä menetelmää pidetään yleisesti halvimpana, pääasiassa siitä syystä, että yhteiskunta voi velvoittaa työnantajat antamaan korvauksetta suoraan tutkimuslaitoksille yleishyödyllisiä tilastotietoja. Viimeksi mainitusta seikasta myös johtuu, että yrityksiltä on yleensä pyritty keräämään totaalinen aineisto. Mikäli todella saadaan tiedot kaikilta kyseisen tuotannon alan yrityksiltä, laskettujen tunnuslukujen luotettavuutta rajoittavat vain mittauksen ja aineiston käsittelyn aiheuttamat virheet.

Teollisuuden työpanoksesta katsotaankin postitse koottavan teollisuustilaston antavan riittävän luotettavat tiedot. Ne voidaan valtaosaltaan perustaa yritysten palkkalistoista saataviin tietoihin tehtyjen työtuntien määristä. Maa- ja metsätaloudessa ja eräillä niihin liittyvillä tuotannon aloilla asianlaita on kuitenkin toinen. Yritysten lukumäärä on erittäin suuri, niiden keskikoko pieni ja niillä käytetään hyvin suuressa määrin yrityksen haltijaperheen omaa työvoimaa. Työpanoksesta pidetään har-

voin kirjaa. Sitä paitsi urakkapalkalla työskentelevän metsätyövoiman työpanosta työnantajan on vaikea mitata, vaikka siihen olisi haluakin. Postitse on tietoja vaikea saada. Yritysten luettelointikin on työlästä.

Vaikkei Suomessa eikä tiettävästi muissakaan maissa ole toistaiseksi pyritty yrityksiltä saamaan tietoja maa- tai metsätalouden koko työpanoksesta, ne tätä tietä voitaneen kerätä tarpeeksi hyvin suunnitelluilla otoksilla. Sama koskee talonrakennusteollisuutta, tienrakennustöitä, uittoa ym. tyypillisiä maaseutuelinkeinoja. Myös työpanoksen kausivaihtelu voidaan arvioida yrityksiltä kerätyllä aineistolla.

Valtakunnan tai sen osa-alueen koko työpanoksen, sen kausivaihtelun ja rakenteen arviointiin yrityksiltä koottu aineisto sen sijaan on vähemmän sovelias. Yrityksiltä kerätyt työvoimatilastot ylen harvoin, tuskin koskaan peittävät koko työvoimaa. Aineiston saaminen sellaiseksi, että se voitaisiin jakaa iän, asuinpaikan tai yhteiskuntaryhmän perusteella alaryhmiin, on vieläkin työläämpää. Valtakunnan työpanoksen kokonaisuusmäärä, sen rakenne ja kausivaihtelu on siten yksinkertaisimmin, halvimmin ja luotettavimmin arvioitavissa väestöstä poimitun otoksen avulla. Myös työpanoksen jakautuminen työvoiman asuma-alueittain tulee tällä menetelmällä luotettavimmin arvioiduksi. Jos sen sijaan on välttämätöntä saada yksityiskohtainen selvitys jonkin elinkeinohaaran työpanoksen jakautumisesta käyttöalueittain, yrityksen suuruusluokittain tai omistajaryhmittäin, on turvaututtava yrityksiltä kerättävään aineistoon.

Myöskin *työttömyyttä* on mahdollisuus tutkia sekä väestöstä otetusta näytteestä mittaamalla että sellaisten laitosten tai järjestöjen välityksellä, jotka ovat yhteydessä työttömien henkilöiden kanssa. Kansainvälinen työtoimisto pitää väestöstä poimittua otosta laitostilastoja parempana työttömyyden tutkimusmenetelmänä. Vuoden 1950 maaseudun työvoiman tutkimuksessa käytetty mittaussuunnitelma antaa myös luppoajasta jonkinlaisen käsityksen. Kun siinä mitataan työttömyysajan määrä, sillä on ainakin mahdollisuus peittää lyhyet avoimen työttömyyden periodit paremminkin kuin USA:ssa ja eräissä muissa maissa käytetyillä työvoiman tutkimuksen menetelmillä.

Jonkin elinkeinohaaran alalla työskennelleiden *henkilöiden kokonaisuusmäärä*, sen jakautuminen väestötunnusten perusteella ja näiden henkilöiden työkauden pituus on arvioitavissa ainoastaan väestöstä otetusta näytteestä. Yrityksen on nimittäin mahdoton seurata siitä vuoden kuluessa poistuneiden työllisten ajankäyttöä.

Mikäli kaikki mainitut tutkimusongelmat on ratkaistava samanaikaisesti, väestöstä otettu näyte on edullisin tietojen lähde, koska se viimeksi mainitun tutkimuskohteen vuoksi on välttämätön ja antaa samalla vain

vähäisin keräyksen lisäkustannuksin tiedot myös muista tutkimuskoh-teista. Kyseisen näytteen mittauksen kohteena on tällöin koehenkilöiden ajankäyttö. Mittausmenetelmistä todettiin vuoden ajankäytön haastat-telu riittävän luotettavaksi ja halvimmaksi menetelmäksi. Sen suurin puutteellisuus on, että alkuvuodenkin työpanostiedot saadaan vasta seu-raavana vuonna. Tämän puutteen poistamiseksi voidaan havaintoaika jakaa osiin, esim. neljännesvuosiin. Kustannukset kuitenkin nousevat mil-tei samassa suhteessa kuin haastattelukerrat lisääntyvät. Jos tutkimuk-sessa halutaan tiedot eri elinkeinohaaroissa (esim. metsätaloudessa) vuo-den aikana työskennelleen työvoiman kokonaisuudesta ja sen raken-teesta, ajankäytön mittaus koko havaintovuodelta on suoritettava samoja koehenkilöitä käyttäen. Tämä osaltaan vaikeuttaa havaintokauden jaka-mista osiin. Jos tästä tutkimuskohteesta sen sijaan luovutaan, on suu-rempi mahdollisuus jakaa havaintokausi, koska kukin osa-aika voidaan tutkia eri otoksesta.

Taloudellisinta on kerätä työvoimatutkimuksen aineisto kantanäyt-teestä (general-purpose sample), joka palvelee yhtäaikaan useita (toistuvia) tutkimuksia, esim. asunto-olojen, tulojen, kulutuksen, verorasituksen ym. arviointeja. Tällaisesta kantanäytteestä kerätään USA:n kuukausittaiset työvoimatiedot, ja sellainen on luotu myös Ruotsissa (DALENIUS 1954 a ja b).

Niin kauan kuin Suomessa ei ole olemassa kantanäytettä ja sen keräily-organisaatiota, sopivin järjestelyin voidaan saada yksiasteinen otos työ-voimatutkimusta varten. Noin 2 500—3 000 koehenkilön yksiasteinen todennäköisyysvalinta määrätyn alimmaisikärajan sivuuttaneesta maa-seudun miespuolisesta väestöstä henkikirjojen perusteella ei tuottane voit-tamattomia vaikeuksia. Onhan kysymys vain keskimäärin noin 5 hen-kilön poimimisesta ja luettelomisesta kunnassa.

Vuoden 1950 maaseudun työvoiman tutkimuksen kaltainen koko valta-kunnan, mutta ei osa-alueitten, miestyövoiman tutkimus Suomessa voi-taisiin siten tehdä niin pienellä aineistolla, että jokavuotinen, jopa vuosi-neljänneksittäin suoritettavan tutkimuksen toimeenpano saattaisi tulla kysymykseen. On ajateltavissa, että vuosineljänneksittäin, kulloinkin eri otoksesta, suoritetaan kuluneen vuosineljänneksen ajankäytön tiedustelu ja kalenterivuoden päätyttyä poimittavasta otoksesta koko edellisen kalen-terivuoden ajankäytön tiedustelu. Täten saataisiin tiedot vuosineljännek-sittäin eri elinkeinoharjojen työpanoksesta sekä työttömyydestä, ja kalen-terivuoden päätyttyä olisi viimeksi kerätystä aineistosta mahdollisuus arvioida myös tärkeimpien elinkeinoharjojen työvoiman kokonaisuus ja sen rakenne.

Merkkien luettelo — *List of symbols*

A. Varianssikaavoissa käytetyt merkit

A. Symbols employed in the variance formulas

Tekstistä selviää, tarkoittaako merkki koko havaintoaikaa (vuotta) vai sen osa-kautta ja koko tutkimusaluetta vai sen osa-aluetta.

It should be clear from the text whether the symbol refers to the total observation period (year) or to its subperiod and whether it refers to the total investigated area or its subarea.

- N miesten lukumäärä perusjoukossa laskettuna kehikon tai väestönlaskennan perus-teella.
total number of males in the population according to the frame or the census.
- n miesten ja lomakkeiden lukumäärä koko saadussa näytteessä.
number of males and forms in the sample collected.
- m miesten lukumäärä näytteen määrättyllä yhteisellä ominaisuudella varustetussa osassa (ositteessa).
number of males with a certain characteristic in a sample subgroup (a stratum of the sample).
- μ miesten lukumäärä m -osittien määrättyllä yhteisellä ominaisuudella varustetussa alaositteessa.
number of males with a certain characteristic in a subgroup of sample subgroup m .

$$F = \frac{n}{N}. \text{ Todellinen otantaosuus.}$$

True sampling ratio.

- f teoreettisesti näyteyksiköiden odotetun lukumäärän perusteella arvioitu otanta-
osuus.
sampling ratio estimated theoretically from the average sample number.

$$K = \frac{1}{F}$$

$$k = \frac{1}{f}$$

$$n' = kn$$

$$m' = Km$$

$$\mu' = K\mu$$

$$\varphi = \frac{m}{n} = \frac{m'}{N}$$

$$\psi = \frac{\mu}{m} = \frac{\mu'}{m'}$$

i näytteen tai sen alaryhmän i :s ajankäyttölomake.
the i th form in the collected sample or its subgroup.

j lomakkeen j :s ajankäyttöryhmä.
the j th activity group of the form.

x_{ij} i :nnen lomakkeen j :nteen ajankäyttöryhmään merkitty muuttuja, joka tarkoittaa päivien lukumäärää.
a variable value (the number of days) in the i th form's j th activity group.

$x_j = \sum x_{ij}$. Kaikkien koko näytteeseen n tai sen alaryhmään kuuluvien lomakkeiden j :nnen ajankäyttöryhmän x_{ij} -muuttujien summa.
Sum of the x_{ij} variables of the j th activity group in the sample n or its subgroup.

\bar{x}_j j :nnen ajankäyttöryhmän x_{ij} -muuttujien keskiarvo koko näytteessä tai sen ositteessa.
arithmetic mean of the x_{ij} variables in the j th activity group of the sample or its subgroup.

$$x'_j = Kx_j$$

z_i i :nteen lomakkeeseen merkitty kyseisen kauden kaikkien itsenäisten x_{ij} -muuttujien summa (päättökimoksessa kaikkien arkipäivien summa).
sum of all independent activity group's x_{ij} variables of the i th form in the period in question (in the Main Survey the sum of the week-days of the period).

$a = z_i$, kun z_i on tai sitä voidaan käytännössä pitää vakiona.
when z_i is or can be taken as a constant in practice.

$z = \sum z_i$. Koko näytteessä tai sen alaryhmässä.
In the total sample or its subgroup.

$$z' = Kz$$

$$p_{ij} = \frac{x_{ij}}{z_i} \text{ . } i\text{:nnessä lomakkeessa.}$$

In the i th form.

$$q_{ij} = 1 - p_{ij}$$

$$\bar{p}_j = \frac{x_j}{z} = \frac{x'_j}{z'}$$

$$\bar{q}_j = 1 - \bar{p}_j$$

s^2 jonkin muuttujan tai estimaatin estimoitu absoluuttinen varianssi. Kyseinen muuttuja tai estimaatti merkitään alaindeksinä, esim. $s_{x_{ij}}^2$.

estimated absolute variance of a variable or an estimate. The variable or estimate in question is indicated with a subscript, e. g. $s_{x_{ij}}^2$.

v^2 jonkin estimaatin estimoitu suhteellinen varianssi (muuttujan tai estimaatin absoluuttinen varianssi jaettuna estimaatin neliöllä, esim. $v_{x'_j}^2 = \frac{s_{x'_j}^2}{x'_j{}^2}$).

estimated relative variance of an estimate (the absolute variance divided by the square of the estimate in question e. g. $v_{x'_j}^2 = \frac{s_{x'_j}^2}{x'_j{}^2}$).

$$c_j^2 = \frac{s_{p_{ij}}^2}{\bar{p}_j \bar{q}_j}, \text{ kun } s_{p_{ij}}^2 \text{ on laskettu kvantitatiivisen varianssianalyysin menetelmin.}$$

when $s_{p_{ij}}^2$ has been computed by quantitative variance analysis.

c_t, c_d usean c_j -estimaatin perusteella pienimmän neliösumman menetelmällä laskettuja tasoitusarvoja.
adjusted values of several c_j estimates computed by the least squares method.

c laskettujen c_t -estimaattien perusteella arvioitu korjauskerroin, jota käytetään arvioitaessa likimäärin estimaattien x'_j ja \bar{p}_j variansseja ja keskivirheitä estimaateista itsestään.

correction coefficient approximated by means of the c_t estimates, used in approximating the standard errors of the x'_j and \bar{p}_j estimates from the estimates themselves.

B. Päättökimpuksen estimaatteja suurennettaessa ja vastausten kadon osuutta arvioitaessa käytetyt merkit

B. Symbols employed in raising and in estimating the non-response ratio in the Main Survey

Tekstistä selviää, tarkoittaako merkki koko maaseutua vai sen osa-alueita.

It should be clear from the text whether the symbol refers to all rural districts or some of them.

T-mies kuukauden 1. tai 15. päivänä vv. 1886—1935 syntyneeksi merkitty miespuolinen henkilö.

T-man *male person whose birthdate is listed as the 1st or 15th of a month in 1886—1935.*

T-lomake päättökimpuksen lomake.

T-form *form used in the Main Survey.*

n_t näytekysiköiden odotettu lukumäärä (täydellinen näyte).
expected sample number.

T_h aineistoon hyväksytyjen T-lomakkeiden lukumäärä.
number of approved T-forms in the sample.

T_a ammattisotilaista täytettyjen ja sen tähden hylättyjen T-lomakkeiden lukumäärä.
number of T-forms completed for men in the regular army and therefore rejected.

T_{50} T-miesten lukumäärä 1. 1. 1950 suoritettuna henkikirjoituksen luettelossa.
number of T-men in the registration list of Jan. 1st 1950.

T_{51} sama kuin edellä, mutta 1. 1. 1951.
the same as the former, but of Jan. 1st 1951.

A_{51} luvun T_{50} perusteella laskettu luvun T_{51} arvio.
estimated value of T_{51} computed from the value T_{50} .

M_{50} kaikkien miespuolisten henkilöiden lukumäärä henkikirjoituksen mukaan 1. 1. 1950.
number of males in the registration list of Jan. 1st 1950.

M_{51} sama kuin edellä, mutta 1. 1. 1951.
the same as the former, but of Jan. 1st 1951.

p_{se} v. 1950 syntyneiden poikalasten ja v. 1950 kuolleiden miespuolisten henkilöiden lukumäärien erotus jaettuna miespuolisten henkilöiden keskiväkiluvulla v. 1950.
births of males minus deaths of males in 1950 divided by the average number of living males in the beginning and at the end of 1950.

p_{ku} v. 1950 kuolleiden 15—64-vuotiaiden miesten lukumäärän suhde kaikkien samanikäisten miesten keskiväkilukuun v. 1950.
deaths of males of 15—64 years in 1950 divided by the average number of living males of the same age in the beginning and at the end of 1950.

p_{la} 31. 12. 1950 laitoshoidossa olleiden miespuolisten 15—64-vuotiaiden henkilöiden lukumäärän suhde samanikäisten miespuolisten henkilöiden koko lukumäärään 1. 1. 1951. Arvioluku.
number of males of 15—64 years under institutional care Dec. 31st 1950 divided by the number of all males of the same age on Jan. 1st 1951. Approximated value.

t_{51} suureen $(1 - p_{la}) T_{51}$ estimaatti.
estimated value of $(1 - p_{la}) T_{51}$.

k suurennustekijä, joka on arvioitu t_{51} -estimaatin perusteella.
raising factor, computed from the estimate t_{51} .

K lopullinen suurennustekijä. Perusjoukossa ja saadussa näytteessä olevien henkilöiden lukumäärien suhde.
final raising factor. The number of males in the population divided by the number of males in the sample.

J_{max} näytteestä puuttuvien koehenkilöiden lukumäärän suhde näytekysiköiden odotettuun lukumäärään. Arvioluku, joka on laskettu t_{51} -estimaatin perusteella.
number of non-responses divided by the expected sample number. Estimate, computed from the estimate t_{51} .

J_{min} sama kuin edellä, mutta laskettu teoreettisen otantaosuuden 24/365.25 perusteella.
the same as the former but computed from the theoretical sampling ratio 24/365.25.

Kirjallisuusluettelo — References

- The agricultural estimating and reporting services of the United States Department of agriculture. 1949. United States Department of agriculture. Miscellaneous publication No 703. Washington, D. C.
- ALAMERI, ROLF — PÖYHÖNEN, PENTTI. 1953. Johdatusta tilastolliseen tutkimukseen. Helsinki.
- Anvisningar till Landstingets näringsutrednings formulär för inventering av samtliga jordbruk i Norrbottens län, avsedda för inventeringsmännen. 1944. Luleå.
- Arbejdsstyrken maj 1951. 1952. Danmarks statistik. Statistiske efterretninger 44, ss. 59—67. København.
- BACHI, R. [1952]. The tendency to round off age returns: measurement and correction. International statistical conferences, December 1951. Bulletin of the International statistical institute. 33, IV, ss. 195—222. Calcutta.
- BERESFORD-PEIRSE, HENRY. 1950. Relation of forestry to agriculture in rural economy. Actes du III^e congrès forestier mondial. Proceedings of the III world forestry congress. 3, ss. 218—222. Helsinki.
- BRUUN, KETIL. 1952. Nuorisotutkimus vuonna 1951. Tilastollisia kuukausitietoja Helsingistä, ss. 193—264. [Helsinki].
- »— 1953. Vanhustutkimus v. 1952. Tilastollisia kuukausitietoja Helsingistä, ss. 73—131. [Helsinki].
- CARLSSON, TORSTEN. 1950. Skogens arbetskraftsproblem. Industriens utredningsinstitut. Industriproblem 1950, ss. 82—110. Stockholm.
- COCHRAN, WILLIAM G. 1953. Sampling techniques. New York — London.
- DALENIUS, TORE. 1949. Om stickprovsundersökningens teknik och metodik. Affärsökonomis skriftserie 25. Stockholm.
- »— 1954 a. Statistiska centralbyråns organisation och verksamhet. V. Utredningsinstitutet. 1: Basurvalet. Statistisk tidskrift, Ny följd 3, ss. 60—78. Stockholm.
- »— 1954 b. Statistiska centralbyråns organisation och verksamhet. V. Utredningsinstitutet. 2: Fältorganisationen. Statistisk tidskrift, Ny följd 4, ss. 177—192. Stockholm.
- DEMING, W. EDWARDS. 1948. Statistical adjustment of data. New York — London.
- DUCOFF, LOUIS J. — BANCROFT, GERTRUDE. 1945. Experiment in the measurement of unpaid family labor in agriculture. Journal of the American statistical association 40, ss. 205—213. Menasha, Wis.
- ELLILÄ, K. J. 1953. Maataloustekniikan kehitys ja maatalouspolitiikka. Kansallis-Osake-Pankki. Taloudellinen katsaus, ss. 251—258. Helsinki.
- Employment, unemployment and labour force statistics. 1948. International labour office. Studies and reports. New series 7 (Part 1). Geneva.
- ENDER, RUDOLF. 1950. Die Forstarbeiterfrage in Österreich forstpolitisch gesehen. Actes du III^e congrès forestier mondial. Proceedings of the III world forestry congress. 3, ss. 304—314. Helsinki.
- Folkräkningen den 31 december 1945 II: 1. 1948. Sveriges officiella statistik. Stockholm.
- Forbruksundersøkelse 1952 etter intervjumetoden. 1953. Norges offisielle statistikk XI. 131. Oslo.
- FOUGSTEDT, GUNNAR. 1953. Maassamuutto Suomessa ja sitä kuvaava tilasto. Kansantaloudellinen aikakauskirja 49, ss. 345—359. Helsinki.
- Fyysillisten henkilöiden verorasitus v. 1949. 1951. Komiteanmietintö N:o 8—1951. Helsinki.
- HANSEN, MORRIS H. — HURWITZ, WILLIAM N. 1946. The problem of non-response in sample surveys. Journal of the American statistical association 41, ss. 517—529. Menasha, Wis.
- HANSEN, MORRIS H. — HURWITZ, WILLIAM N. — MADOW, WILLIAM G. 1953. Sample survey methods and theory. I. New York—London.
- HEIKINHEIMO, LAURI. 1950 a. Maaseudun työvoimavarojen arvioinnin tarpeellisuus. Metsätaloudellinen aikakauslehti 67, ss. 81—84. Helsinki.
- »— 1950 b. Tarkistamattomia ennakkotietoja Helsingin yliopiston metsätaloustieteellisen laitoksen metsätyöläistutkimuksista vv. 1949/50. Mimeogr.
- »— 1951 a. Työvoimaterminologiaa. Maatalous 44, ss. 233—234. Helsinki.
- »— 1951 b. Forest work wage statistics. Seminar on statistical techniques. Lectures on forestry statistics, ss. 36—42. Helsinki. Mimeogr.
- »— 1954 a. Suomen metsätyöntekijä. The Finnish forest worker. Suomen puun miehiä. Finnish timber men, ss. 98—109. Helsinki.
- »— 1954 b. Maaseudun miestyövoiman arkiajan käyttö. Maaseudun työvoiman tutkimuksia II. Summary: Use of rural manpower in Finland. Finnish rural labour force studies II. Julkaistava konekirjoite.
- HELANDER, A. BENJ. 1923. Metsätalouden tarjoamasta ansiotyöstä. I. Referat: Der Arbeitsbedarf in der Forstwirtschaft. Acta forestalia fennica 26. Helsinki.
- HIESS, FRANZ. 1931. Methodik der Volkszählungen. Jena.
- HOSEMAN, H. 1950. Bestehen solare und lunare Einflüsse auf die Nativität und den Menstruationszyklus? Zeitschrift für Geburtshilfe und Gynäkologie 133, ss. 263—285. Stuttgart.
- HYPPÖLÄ, JORMA — TUNKELO, AARRE — TÖRNQVIST, LEO. 1949. Suomen väestöä, sen uusiutumista ja tulevaa kehitystä koskevia laskelmia. Tilastollisia tiedonantoja, julkaissut Tilastollinen päätoimisto. 38. Helsinki.
- HYRENIUS, HANNES. 1947. Skogsbrukets arbetskraft. Karlstad.
- IHATSU, SIRKKA. 1953. Kontiolahden kunnan Kunnasniemen ja Jakokosken kylien sekä Lehmon kansakoulupiirin miestyövoiman ajankäyttö. Konekirjoite.
- JAFFE, A. J. — STEWART, CHARLES D. 1951. Manpower resources and utilization. New York—London.
- JÄNTTI, AUGUST. 1953. Maataloustuotannon teknillinen kehittäminen ja viljelystilan koko. Maatalous 46, ss. 139—141. Helsinki.
- KAHMA, JAAKKO. 1953. Tuotantoreservit käyttöön — teollisen toiminnan edellytyksiä vahvistamalla. Unitas 25, ss. 119—121. Helsinki.

- KANTOLA, MIKKO. 1949. Tutkimuksia ammattitaidon ja työkalujen merkityksestä metsänhakuutöissä. Summary: Investigations into the significance of professional skill and tools in logging. Työtehoseuran julkaisuja 52. Helsinki.
- KEKKONEN, URHO. 1952. Onko maallamme malttia vaurastua? Toinen painos. Helsinki.
- KEYNES, JOHN MAYNARD. 1951. Työllisyys, korko ja raha. Porvoo — Helsinki.
- KLIMEK, HENRYK. 1950. Employment and unemployment in Polish forest economy. Actes du III^e congrès forestier mondial. Proceedings of the III world forestry congress. 3, ss. 280—283. Helsinki.
- KOROLEFF, A. et al. 1951. Stability as a factor in efficient forest management. Pulp and paper research institute of Canada. Woodlands research index 89. Montreal.
- LAKARI, O. J. 1934. Valtionmetsien työtarjonnasta. Referat: Über das Arbeitsangebot in den Staatswäldern. Acta forestalia fennica 40. Helsinki.
- Laskijan opas. Yleinen väestölaskenta joulukuun 31 päivänä 1950. [1950]. Helsinki.
- LINDFORS, JARL. 1937. Metsätaloutemme työn tarpeesta. Suomen paperi- ja puutavara-lehti. Erikaisnumero 7 A, ss. 317—318, 320—322, 324. Helsinki.
- LINNOILA, B. W. [1939]. Metsätyöväen asunto-olot hakkuu- ja ajokautena 1938—1939. Sosiaalinen aikakauskirja 33, ss. 484—492. Helsinki.
- LÄHDE, ARVO. 1940. Metsä- ja uittotyöläisten asunto- ja ravinto-oloista. Referat: Über die Wohnungs- und Ernährungsverhältnisse der Wald- und Flossarbeiter. Silva fennica 51. Helsinki.
- Maaseudun elinkeinokomitean mietintö. Komiteamietintö 1951. Helsinki. Mimeogr.
- MACDONALD, D. A. 1950. Employment and unemployment in forestry in Canada. Actes du III^e congrès forestier mondial. Proceedings of the III world forestry congress. 3, ss. 287—303. Helsinki.
- MAKKONEN, Olli. 1950. Hakkuutöiden aikatutkimustulosten soveltaminen käytäntöön. Summary: Practical application of the results of time studies in logging. Metsäteho. Julkaisu 25. Helsinki.
- Methods of family living studies. 1949. International labour office. Studies and reports. New series 17. Geneva.
- Metsätilasto. [Eri vuosilta]. Suomen virallinen tilasto XVII. Helsinki.
- NIITAMO, OLAVI. [1954]. Työn tuottavuuden kehitys Suomen teollisuudessa vuosina 1925—1950. [Helsinki]. Mimeogr.
- OSARA, N. A. 1950. Relationship of forestry to agriculture in rural economy. Actes du III^e congrès forestier mondial. Proceedings of the III world forestry congress. 3, ss. 243—245. Helsinki.
- »— 1953. Eräitä valtion metsätalouden ongelmia ja tavoitteita. Kansantaloudellinen aikakauskirja 49, ss. 431—450. Helsinki.
- PELTARI, P. J. 1935. Yksityismetsien työtarjonnasta. Referat: Über das Arbeitsangebot in den Privatwäldern. Silva fennica 35. Helsinki.
- PIHKALA, K. U. 1947. Työn käyttö suomalaisessa maataloudessa. Asutuskysymys ja maataloutemme kehitysmahdollisuudet, ss. 92—107. Helsinki.
- »— 1952. Ihmistyön käyttö Suomen maatalousviljelmillä v. 1950. Maatalous 45, ss. 167—173. Helsinki.
- Pohjoismainen tilastosanasto. Nordisk statistisk nomenklatur. 1954. København.
- The preparation of sampling survey reports. 1950. Statistical office of the United Nations. Statistical papers. Series C, 1. Lake Success, N. Y.

- Puun maakuljetusolojen kehittäminen. 1952. Vuoden 1949 metsätalouskomitean mietintö N:o 1. Silva fennica 73. Helsinki.
- PÖNTYNEN, V. 1936. Metsän hakkuun ja ajon sekä puutavaran uiton työn kysynnästä. Referat: Über die Arbeitsnachfrage bei Abtriebs- und Abfuhrarbeiten sowie in der Holzflösserei. Acta forestalia fennica 42. Helsinki.
- SAARI, EINO. 1922. Kotitarvepuun kulutus maaseudulla. Helsinki.
- »— 1928. Metsä- ja uittotyöväen kysymys. Maa ja metsä IV, Metsätalous I, ss. 278—292. Porvoo.
- SALONEN, K. D. J. 1953. Piirteitä maaseudun työttömyydestä. Kansamme talous, ss. 59, 62. Helsinki.
- SAUER, WALTER. 1929. Die Waldarbeiterverhältnisse in den württembergischen Staatsforsten. Marbach a. Neckar.
- SAULI, LIISA. 1951. Viljelijäperheiden elintaso maataloushallituksen kirjanpitotiloilla tilivuosina 1935/36—1948/49. Summary: The standard of living of farm families on the bookkeeping farms of the board of agriculture in the fiscal years 1935/36—1948/49. Acta agralia fennica 75. Helsinki.
- SAURIO, ELLI. 1947. Maalaisemännän ajankäyttö suhteessa talouden laatuun ja henkilöarakeenteeseen. Summary: Use of time of rural Finnish homemakers in relation to the character and composition of the household. Helsinki.
- SIPILÄ, ARVO. 1938. Suomen työoikeuden käsite ja järjestelmä sekä suhde sosiaaliseen lainsäädäntöön. Helsinki.
- Skogsarbetarnas levnads- och arbetsförhållanden i Värmland, Dalarna och Norrland. 1916. Sveriges officiella statistik. Socialstatistik. Stockholm.
- Skogsbygdens arbets- och levnadsvillkor. I. 1938. Sveriges officiella statistik. Socialstatistik. Stockholm.
- SNELLMAN, G. R. 1914. Undersökning angående sågindustrin samt därmed i sammanhang stående afverkning, flottning och inlastning i Finland. Arbetsstatistik XVI. Helsingfors.
- STEINBERG, JOSEPH. 1953. The current population survey. HANSEN, MORRIS. H. — HURWITZ, WILLIAM N. — MADOW, WILLIAM G.: Sample survey methods and theory. I, ss. 559—582. New York—London.
- Suomen tilastollinen vuosikirja. [Eri vuosilta]. Helsinki.
- SÄLLFORS, TARRAS. 1945. Teollisuuden työntutkimukset. Helsinki.
- Tapaturmatilastoa. [Eri vuosilta]. Suomen virallinen tilasto XXVI A, uusi sarja. Helsinki.
- Teollistamiskomitean mietintö. 1951. Komiteamietintö N:o 12—1951. Helsinki.
- TREITSCHKE, KURT. 1928. Die wirtschaftlichen und sozialen Verhältnisse der preussischen Staatsforstarbeiter. Schriften des Deutschen Landarbeiter-Verbandes 21. Berlin.
- Tutkimus metsä- ja uittotyöntekijäin oloista keväällä 1921. 1923. Suomen virallinen tilasto XXXII. Sosiaalisia erikoistutkimuksia 4. Helsinki.
- Työttömyyskomitean 1950 mietintö. [1950]. [Helsinki]. Mimeogr.
- Uittotilastoa. [Eri vuosilta]. Aineiston kerännyt Suomen uittajainyhdistys, r.y. Helsinki.
- WALLANDER, JAN. 1948. Flykten från skogsbygden. With an English summary. Stockholm.

- Valtakunnansuunnittelukomitean mietintö. 1954. Komiteamietintö N:o 4—1954. (I osa). Helsinki.
- Valtiovarainministeriön kansantalousosaston suhdannesarjat. 1950, 1951. Helsinki.
- WEATHERBURN, C. E. 1947. A first course in mathematical statistics. Cambridge.
- WEIGER, EBERHARD. 1954. Die Arbeiterfrage in der deutschen Forstwirtschaft. 2. erweiterte Auflage. München.
- WEISS, SAMUEL — WILLIAMS, FAITH M. [1952]. A program of labor statistics. International statistical conferences, December 1951. Bulletin of the International statistical institute. 33, IV, ss. 265—280. Calcutta.
- W[ICHMANN], A[SSAR]. 1937. Metsä- ja uittotyöläistilastoa vuodelta 1936. Suomen paperi- ja puutavaralehti 19, ss. 500, 502, 504. Helsinki.
- WICHMANN, ASSAR. 1939. Metsä- ja uittotyöläiset sekä heidän työansionsa. Suomen paperi- ja puutavaralehti 21, ss. 380, 382, 384. Helsinki.
- WIESLANDER, G. 1948. Arbetskraftsbalansen inom Norrbottens läns skogsbruk intill år 1970. Umeå.
- WUORIO, JORMA. 1952. Työntekijän käsitteestä työoikeuden eri aloilla. Lakimies 50, ss. 690—721. Helsinki.
- VUORISTO, ILMARI. 1934. Perä-Pohjolan metsätyöläisten ikä-, perhe- ja maanomistussuhteista. Metsälehti N:o 11. Helsinki.
- Väestölaskenta 1950, ohjeet D₁ ja E sekä lisäohje D₁:een. [1951]. Mimeogr.
- Väestötilastoa. 1951. Suomen virallinen tilasto VI A 105. Helsinki.
- Väestötilastoa. 1952. Suomen virallinen tilasto VI A 106. Helsinki.
- YATES, FRANK. 1949. Sampling methods for censuses and surveys. London.
- Yleinen maatalouslaskenta v. 1941. I. 1945. Suomen virallinen tilasto III. Maatalous 38: 1. Helsinki.
- Yleinen maatalouslaskenta v. 1950. I. 1954. Suomen virallinen tilasto III. Maatalous 45: 1. Helsinki.
- ØDEGÅRD, JOHAN — RESEN-MANDT, M. [1939]. Skogsarbeidernes lønns- og levekår. Kongsvinger.

Method of surveying forest labour

Finnish rural labour force studies I

Summary

1. Introduction

11. Terminology

According to the definition of the International Labour Office (Employment . . . 1948, p. 119) the labour force on a given day includes all the persons employed on that day, called *workers* in the following, and *the unemployed*.

Forest work refers to manual work in forestry, and the *forest workers* on a given day refer to all the persons engaged mainly in forest work on that day. The term *forest labour force* in the following covers the forest workers only. *The forest labour force of a given month or year is taken to include every person who during that month or year has been engaged in forest work at least one day*. Thus defined, the forest labour force of a year in practice exceeds that of any month or day of the year in question.

According to the independence of their work, forest workers are divided into *paid forest workers*, working on the basis of a labour contract, and *unpaid forest workers* who do forest work for their own account or without pay for the account of some other member of the family.

The forest workers of a year are divided, according to the importance forest work has for their sustenance, into *professional* and *seasonal* forest workers. For the former forest work is the principal, for the latter a subsidiary source of income. A similar classification is applied to the paid forest workers.

By the permanence of the employment, all forest workers in theory, but only professional paid forest workers in practice are classified as *permanent* or *occasional*. The *permanent* (professional and paid) *forest worker* is supposed to be engaged in forest work for one and the same forest owner throughout a sequence of years, just like a factory worker.

By the nature of their work, the workers engaged in the *logging of wood for sale* and in the *logging of wood for domestic use* are distinguished from the rest of the forest labour force. In both groups the *fellers* and the *haulers* may also be distinguished.

As forestry and floating in Finland are closely associated, the *forest and floating labour force* is often dealt with in the following as a combined group, applying the classification outlined above. The labour force engaged in agriculture and other seasonal

rural occupations is also defined and classified on these grounds. It is to be noted that the male labour force employed in agriculture and forestry in Finland in a given year largely overlap if thus defined.

Some investigators measure the size of the labour force employed in a certain occupation by working-time units (e.g. KOROLEFF 1951, p. 11). In that case, however, it would be clearer to replace the term »labour force» by the concept »labour input» (cf. JAFFE & STEWART 1951, p. 205). The *input of labour to forestry*, in terms of days, refers to the total of working days employed on forestry during the period involved. The (total) *labour input of the forest labour force* is the sum of the time units (days) spent on *all kinds of work* by the forest labour force during a given period. The *forest labour input of the forest labour force* is the sum of days spent by the same group of persons on *forest work*. Accordingly, e.g. we have the terms »(total) labour input of a forest worker» and »forest labour input of a forest worker».

The labour input on the worker's own account or without pay for the account of another family member is termed in brief the *input of unpaid labour*. Accordingly, the labour input of a worker against wages is called the *input of paid labour*.

Week-days in the following refer to all days except the holidays and other official days off not included in labour input. The *week-days* included in labour input are called *working days*, their sum the *working time*. The utilization of week-days (the *round of activity*) by a person or group of persons is divided into two main *activity groups*: *labour input week-days* and *non-working week-days*, such as those spent on illness, holiday, studies, unemployment etc. The use of the male labour force's week-days during a year (the annual round of activity) is termed in brief »*use of manpower*».

12. Earlier methods of forest labour force studies

Efforts have been made to study the amount of forestry and floating labour input, seasonal fluctuation and future development and the amount, structure and employment of the paid labour force engaged in forestry and floating, principally by complete enumerations based on data released by the employers (HELANDER 1923, TREITSCHKE 1928, SAUER 1929, LAKARI 1934, LINDFORS 1937, WICHMANN 1937 and 1939, SKOGS-BYGDENS ... 1938, WIESLANDER 1948, WEIGER 1954 and the population censuses of various countries). In some cases the partial survey method has been employed (PÖNTYNEN 1936, the official Finnish forest labour force statistics). Investigations into the social conditions of the paid forest and floating labour force have mainly been based on partial, purposive or accidental selected samples taken from the work-sites (SKOGS-ARBETARNAS ... 1916, TUTKIMUS ... 1923, WUORISTO 1934, SKOGSBYGDENS ... 1938, LINNOILA 1939, KANTOLA 1949). Sometimes also a complete enumeration has been applied by means of inquiries submitted to the employers (SNELLMAN 1914) or local experts (SKOGSARBETARNAS ... 1916), less often by systematic selection (WALLANDER 1948). In the last-mentioned and a few other cases (ØDEGÅRD & RESEN-MANDT 1938, CARLSSON 1950) the sampling unit has been the individual worker and not the work-site.

13. Random sample in the survey of the forest and floating labour force

Labour force enumerations meant to be complete, in which information is collected from employers or other institutions, seldom cover the total labour force (JAFFE & STEWART 1951, p. 49). It is particularly difficult to try and include all forest work-sites in the statistics as their listing involves almost insuperable difficulties. To obtain a sample acceptable to statistical theory is also laborious. In addition, a work-site is not nearly always the best possible sample unit to represent the forest and floating labour force. The structure (and type) of the labour force varies seasonally at the work-sites, and thus a sample taken from work-sites at a given moment may give a highly biased picture even of the structure of the labour force. It is preferable to take a labour force sample from the population, using the individual as the sampling unit without work-sites as the first-stage sampling units even.

The limitation and listing of the population of such a sample is another difficulty. In Scandinavian conditions, however, urban dwellers, all women and very young boys can be excluded from the population. Such a high proportion (approx. 70 %) of the remaining population in Finland belongs to the forestry and floating labour force that the simplest procedure is to take a forest and floating labour force sample from it. (There exists a dependable, annually revised list of all inhabitants — registration list —, but picking out from this list the persons constituting the whole forest and floating labour force is an insurmountable task.) This gives a sample at the same time of the male rural labour force which, due to the interconnection of the labour force problems of the various rural occupations, is essential to a thorough study of forest and floating labour in the national economy sense. At the same time a subgroup of the sample gives us a dependable sample for the investigation of the social conditions of paid forest and floating workers.

As a sample of the total rural male population of working age is taken for the present study of the forest and floating labour force, the investigation assumes the character of a general labour force survey and is thus similar in principle to the labour force sample surveys effected in the USA and certain other countries.

14. Determination of the round of activity in labour force survey

The observation method applied in USA labour force surveys is a kind of incomplete weekly round of activity interview. The individual interviewed is asked to state his total labour input in hours during the observation week and his main occupation (JAFFE & STEWART, pp. 40—43). On the basis of these data the labour force is divided into workers employed in the various occupations and the unemployed. No efforts are made to ascertain the total labour input of the labour force and its structure. Experience gained in Japan and Puerto Rico of this observation method indicates that the practical value of the statistics obtained is very small (l. c., pp. 469—476). For countries where underemployment is a characteristic phenomenon the investigators recommend employment of the labour input concept in labour force surveys (l. c., p. 54). This is attained by distributing the respondent's round of activity during the observation period into

activity groups. From these data it is possible to calculate the total labour input of the labour force and its structure, including the sum of unemployment time, leisure hours etc. Division of the total output by the corresponding labour input gives a figure indicating productivity from which it is possible to estimate underemployment. An attempt to estimate the latter can also be made by putting the question direct to the respondent.

Round of activity observation methods have been developed primarily in the science of management, sociology and home economics. The methods most frequently used are the direct observation of time expenditure, the interview and the bookkeeping methods. The interview method is most appropriate in the case where a large sample is to be taken. The observation period in the round of activity interview is usually short, though recently observation periods as long as one year have been adopted, as for family living studies by the interview method. In the Finnish agricultural censuses of 1941 and 1950 a one-year observation period was successfully tried out in the inquiry into farm labour input (Yleinen maatalouslaskenta ... 1945 and 1954). In Norrbotten, Sweden, an interview to find out the annual round of activity of some 30 000 persons was effected in 1944. It involved a complete enumeration (WIESLANDER 1948). The unit of measurement in these inquiries was the day, and only the most important activity groups were aimed at.

In the above surveys the annual round of activity was studied retrospectively by a single interview. To reduce response errors, the observation period may be split up e.g. into months, and an interview effected at the end of the subperiod. It is easiest to use a different sample for each subperiod. It is also possible to exclude a number of the subperiods and estimate the results for these intervals by interpolation. The cheapest method, however, is the one in which the manpower's annual round of activity is investigated by a single interview at the end of the year. Its great advantage is that it provides information on the time budget of identical sampling units; in this way the size and structure of e.g. the national forest and floating labour force can be ascertained, a result of great importance from forestry point of view.

2. Pilot surveys

21. Organization of pilot surveys

211. Pilot Survey I. Interview surveys in the communes of Ranua and Kontiolahti in 1949.

In 1949 the present author was commissioned to plan a survey of the conditions of forest workers in Finland. After the conclusions outlined above had been reached, it was decided to start pilot surveys in one commune of Lapland — Ranua, and one of East Finland — Kontiolahti. Their purpose was to try out various observation methods of discovering the annual round of activity of men and to test the sampling methods to be applied. To start with, an experiment was made on the lines of the interview method applied in Norrbotten, Sweden (WIESLANDER 1948). The main form used in

this survey, Pilot Survey I, is reproduced in the appendices. With minor exceptions, it is identical with the form employed in Norrbotten.

In the sample communes a random sample was taken from males born in 1883—1931 entered in the registration list. The population was stratified by villages, age classes and marital status. The size of the population, number of sampling units and sampling ratios are shown in Table 1/p. 29. The interviews were effected by the author plus 5 field-workers in July-August 1949. The material was treated by adding together the corresponding variables of the forms by villages and within them by social groups. The distribution of the sample of Pilot Survey I into social groups is given in Table 2/p. 30. The treatment of the material was an endeavour to study the factors affecting the dispersion of labour input and unemployment time in order to clarify the stratification possibilities and size of sample needed. On the other hand it was desired, by comparing the results with other available data, to ascertain any possible bias in the responses. Response bias was also studied by comparing the data obtained by Pilot Survey I on the input of paid labour of the males in the sample with the entries made by employers in the tax books¹ of the same persons.

212. Pilot Survey II. Bookkeeping survey at Ranua in 1950.

In order to study the possibilities of round of activity bookkeeping, the fluctuations in the structure of the round of activity, and the response bias on interview, a 10 % random sample (133 persons) was taken from the male inhabitants of Ranua commune born in 1884—1933, using the registration list as a frame. During 1950 they were given a special bookkeeping calendar at the beginning of each month. A section of a completed calendar is reproduced in the appendices. The respondents were asked to indicate daily with a cross the principal activity of that day. A special network of collectors was organized in the commune for Pilot Survey II. They were entrusted with distributing and collecting the calendars and with supplementing by interview any calendars unsatisfactorily completed. All the calendars issued were in fact recovered completed although the respondents and collectors received no remuneration for their work. The achievement of this result, however, required voluminous correspondence and numerous visits to the collection area by the survey staff.

213. Pilot Survey III. Experiment with interview forms in 1950.

In the spring of 1950 an experiment with interview forms of four different types was carried out in the neighbourhood of Helsinki. Only 48 forms in all were completed in this rapid survey.

¹ The books in which all wage payments, advance tax withholdings and the periods covered by them are entered, comparable with the PAYE records kept in Britain — Translator.

22. Results of the pilot surveys

221. Sampling method

In Pilot Survey I it was fairly easy to reach all the respondents in the villages with a sampling ratio < 1 . By contrast, in the area projected for complete enumeration, contacting all the persons entered in the list would have involved heavy additional cost. This was mainly due to the fact that a sample list is of great interest to the population, and hence most of them try to help the collector locate the few persons involved. A total list arouses little interest and calls for numerous trips to and fro.

The dispersion of the variables indicating the amount of labour input proved to be great within both the commune and village strata, and similarly within both the age group and the social group strata, compared with the dispersions between these strata. Such stratification does not appreciably increase confidence in the estimates. By the occupations entered in the registration list, however, the men may be divided into two strata: (1) all white-collar and factory workers, and (2) other workers. In most cases variation within these strata probably is less than between the strata.

As a dependable registration list of the population is available in Finland it is fairly easy to take a simple random sample from the population for a labour force survey. Provided collection expenses can be made approximately proportional to the number of sampling units, this is the cheapest possible sampling method and simplifies greatly the estimation of the confidence of results.

222. Size of the simple random sample

It was possible from the materials of Pilot Surveys I and II to estimate the size of sample required for a certain confidence requirement of the estimates. The calculations refer to the size of simple random sample selected from an infinite population. The variance analysis of the present investigation employs symbols based on a principle similar to that used in the manual by HANSEN, HURWITZ & MADOW (1953). They are listed on pp. 137–141.

As, in Pilot Survey I, $z_i = \text{constant } a$, $v_{\bar{x}_j} = v_{\bar{p}_j}$ and is obtained from formula 1/p. 37 by substituting n_1 for m_1 , and n_2 for m_2 . Accordingly, the size of the simple random sample required for a certain confidence is obtained from formula 2/p. 37 after the relative variances of the estimates, $v_{\bar{x}_j}^2 = v_{\bar{p}_j}^2$, have been calculated from the pilot survey. For Table 3/p. 39 three series of n_2 estimates have been calculated from Pilot Survey I, assuming that the confidence requirement is $v_{\bar{x}_j} = v_{\bar{p}_j} = \pm 0.1$. These $n_{(v = 0.1)}$ estimates are entered in Fig. 1/p. 40 as functions of the \bar{p}_j estimates and adjusted ocularly. The result shows that a sample of 1 000–2 000 males will ensure satisfactory confidence provided the homogeneity of the population is roughly identical with that of the communes of the pilot survey.

223. Observation method

In a yearly round interview survey the responses in individual cases may be expected to differ appreciably from the true figures. However, response variance is easy to handle. The gravest threat to the reliability of the results is response bias (HANSEN et al. 1953, pp. 89–91).

The response bias of the pilot surveys was studied as follows:

- (1) by comparing the results of Pilot Survey I with tax book material,
- (2) by getting the interviewers of Pilot Survey I to make subjective estimates of the accuracy of the responses,
- (3) by applying common sense standards to the results obtained.

The results of comparison (1) are given in Tables 4/p. 43 and 5/p. 44. Even the tax book data on the amount of the input of paid labour, here serving as the object of comparison, may differ one way or the other from the true input of paid labour. For this reason the comparisons are only a pointer to show that there was nothing indicative of systematic overstating or understating by the interviewed persons of their labour input. The deviations in the figures compared are no greater at the beginning than at the end of the observation period.

The collectors of the material of Pilot Survey I classified 362 interviews as follows: 37 % were »accurate», 60 % »fairly accurate» and 3 % »inaccurate». However, this result too cannot be considered strictly as evidence.

The interviewed were also classified into occupational groups in the pilot survey on the basis of the ease of response; the order of the groups, from easiest to most difficult, was:

- (1) White-collar workers, permanently employed manual workers, and all unpaid workers, excluding farmers.
- (2) Farmers who did not undertake paid work for others.
- (3) Occasional workers engaged in paid work only.
- (4) Men engaged in farming for their own account and in occasional paid work.

With few exceptions, the persons interviewed in Ranua commune came in Group (4). The data obtained from Pilot Surveys I and II on the amount, structure and seasonal fluctuations of labour input and their correlation with the social group were compared with other available data and judged by common sense standards (HEIKINHEIMO 1950 b and Työttömyyskomitean ... 1950). Nothing indicative of bias in the results was observed. An interesting result of Pilot Survey II is shown in Table 6/p. 48 and Fig. 2/p. 50.

Although Pilot Surveys I and II provided no conclusive evidence as to the accuracy of the interview method, the evidence obtained, taken in conjunction with the experience gained in Norrbotten (WIESLANDER 1948, p. 9), reinforced confidence in this observation method to such an extent that it was decided to adopt it for the Main Survey also. The best study of response bias, a comparison of interview and bookkeeping surveys, could only be effected after completion of the Main Survey. Its results are reported in Chapter 332.

Although Pilot Survey II (bookkeeping) turned out well it was so laborious and expensive that it proved impossible to apply in a sample of several thousands of persons.

224. Interview form

The pilot surveys showed that the best procedure was to divide the observation period into calendar months. The best unit of measurement, simplest from the practical point of view, was the week-day. The total (a) of all week-days was to be printed in the blank form. In order to allow for the holidays and other official days off to be included in the labour input, it had to be possible to make this total (a) greater if necessary, in which case it also became a variable (z_i).

In Norrland and in Pilot Survey I the unpaid labour input was reduced by the estimated number of underemployed days. This method proved erroneous. The labour input in the present instance was to be entered on the form exactly as indicated by the respondent, but the number of the underemployed days included was to be entered as a separate activity group. The input of labour to relief work was to be entered in the same way. The use of underemployed days as ascertained by Pilot Survey II is shown in Table 7/p. 53.

3. Main Survey. The 1950 Rural Labour Force Survey

31. Purpose and limits of the Main Survey

A general population census was taken in Finland at the turn of 1950. As material for other investigations also could be collected in this connection at relatively low expense a committee appointed for the purpose resolved that a sample was to be taken in connection with the census for a survey of the forest labour force and, at the same time, of the total rural manpower. The survey, carried out by the present author with co-workers in the Helsinki University Institute of Forest Economy (Chief: Professor EINO SAARI), was entitled «The 1950 rural labour force survey». In the following it will be called briefly the Main Survey.

The purpose of the Main Survey was threefold: (1) to ascertain the total labour input of the male rural labour force and its subgroups in 1950, concentrating on the forest labour input, seasonal fluctuations in the labour input, and the dependence of forest labour input on the labour input of other occupations and their seasonal fluctuations; (2) to study the amount and seasonal fluctuations of the unemployment time of this labour force and its subgroups; and (3) to establish the total forest labour force in terms of persons employed in 1950, to study its social, occupational and age structure and social conditions.

The survey was limited to the male rural population of 15—64 years, excluding persons confined to institutions and the regular army. After the Main Survey and the census had been completed it was found that a better result would have been attained if males over 64 also had been included in the sample (Table 8/p. 57). As 33 % of the boys of 15 proved to be employed in forest and floating work, 14-year olds at least should also have been included in the sample.

32. Main Survey method

321. Collection and size of the material

Efforts were made to select from the population for inclusion in the sample all males born on the 1st or 15th of any month in 1886—1935. In the following they are termed the T-men. These sampling units were selected in connection with the census by the enumerators. Thus the sampling method was systematical. This method was favoured because the census was taken on the 31st day of December 1950, on which day the registration list available, brought up to date for the preceding January 1st, was a year old. It was desired to include in the sample the T-men who had migrated into the communes in the course of 1950. Sampling by random numbers from the frame would also have been difficult to organize technically. In addition, a desired requirement was to compute the advance results of the Main Survey before information on the size of the population became available from the census.

HOSEMANN (1950) has shown that no periodicity of approx. 15 or 30 days exists in nativity. Thus the sample selected in the way outlined, consisting of sampling units *actually born* on the 1st or 15th day of a month, may be considered to correspond to a simple random sample.

For each T-man of the population, the census enumerator was to complete a so-called T-form, a completed specimen of which with English translations of the main text is reproduced in the appendices. Its most important section was the annual round of activity table on pp. 2 and 3 of the form.

A total of 47 895 forms were returned to Helsinki University; of this number, however, 1 023 were rejected as not belonging to the sample and 19 forms for other reasons. The final size of the sample obtained therefore was 46 853 males. Their distribution by counties is shown on p. 63.

322. Computational procedure

The initial material was subjected to a three-phase check. In the first phase, 7 persons checked according to a certain formula (reprinted in Finnish in the appendices) that the forms had been correctly and fully completed. Incomplete and indistinct forms (some 15 %) were transferred to the next phase of checking and returned to the local authorities for supplementary information. In the third phase two persons rechecked the entire material. After these checks the forms were codified and the data transferred to punched cards (the IBM method). Two punched card materials were made. For each form a «person card» was punched for Codes 1—27 (in the right-hand margin of pp. 1 and 3 of the form). In addition a «time card» was punched for each completed activity group on p. 3 of the form by the mark sensing method. On the time cards both the punchings of the person cards and the mark sensing entries were punched by the mark sensing machine. The person card of a form and a time card are reproduced in the appendices. The person card material (46 853 cards) was treated by a sorter counter. This material was used for the investigations into the structure of the labour force, with the individual as the calculation unit. The time card material (some 300 000

cards) was divided by the sorter into two groups: those unable to work and the able. In the latter group forest workers were treated separately. The figures of each group were added together by a tabulator and reproducing summary punch for each commune and activity group, raised for each commune by means of a multiplying punch and finally tabulated by districts and for the whole country. All the control couplings available were used.

323. Raising method

3231. Raising without knowing the size of population

The raising factors were preliminarily estimated, before the age class tables of the census were completed, by communes from the formula $k = 365.25 t_{51} / 24 (T_h + T_a)$, in which the size of a complete sample, t_{51} , was estimated from Formula 4/p. 69. The explanation of the symbols is given on pp. 140—141. The main results of the survey were computed as advance information with the raising factors estimated in this way.

3232. Final raising method

When the census information on the true size (N) of the population by communes was ready it could be compared with the estimated population (n'). The result of the comparison is shown by counties in Table 9/p. 71. On the basis of these data the estimates were corrected by counties with the coefficient N/n' . The final raising factors and sampling ratios are given in Table 10/p. 72.

3233. Accumulation of birth dates

To find the factor responsible for the considerable differences between the figures N and n' , varying in the different parts of the country, the National Pensions Institute investigated from its punched cards index of the total adult population of the country whether the ratio of cards with birthdays occurring on the 1st or 15th was different from the expectation value $24/365.25$. The investigation covered a total of 407 767 men and women born in 1890—1935 in various parts of the country. 28 631 persons were found to have birthday entries of 1 or 15. This exceeds the expectation value by 6.87 ± 0.19 % (simple standard error). Recorded births on the 1st or 15th were computed from 332 877 cards, separately for each group. Recorded births on the 1st of the month exceeded the expectation value by 9.72 ± 0.97 %, on the 15th by 3.26 ± 1.93 %. The number of persons entered as born on the 1st exceeds the expectation value highly significantly. The difference is small or nil in urban areas and greatest apparently in the eastern and northern parts of the country.

This accumulation phenomenon of birth dates, to which the erroneousness of the raising calculations also was due, can hardly be explained otherwise than by assuming that the inhabitants, particularly in the least developed areas of the country, have a tendency to round off their birth dates, especially with the figure 1 which is easy to

remember. It is probable that similar mistakes occur in connection with transcribing the registration lists of the population. The phenomenon is identical in character with the accumulation of age classes observed e.g. by HIESS (1931, pp. 51—64) and BACH (1952) in the censuses of several countries. As in Finland sampling by birth dates has often been adopted in preference to sampling by random numbers, this accumulation phenomenon, unless it is allowed for, may give rise to biases difficult to trace in the sample.

33. Accuracy of the Main Survey

331. Sampling biases

The non-response ratio of the main survey has been computed in Table 11/p. 77 by counties. Two methods were used: J_{max} and J_{min} from the formulas on p. 76. It is difficult to decide which of the estimates provides the most accurate result as the governing factor is the number of males in the sample whose true birth dates were not the 1st or 15th. In any case, the non-response ratio is fairly high.

It was impossible in the Main Survey to obtain a sample of the missing responses and thus investigate the representativeness of the material collected. The possible bias of the sample due to non-response was studied by comparing the age distribution and occupational distribution of the population obtained from the census with the corresponding distributions of the sample. As the relevant distributions of the true population could not be computed, the object of comparison is called the «population» (Pe). It includes e.g. persons in institutions, a group not included in the sample.

Table 12/p. 79 and Fig. 3/p. 80 show the comparison of age distributions by counties. The uniformity must be considered satisfactory. To save expense, occupational distributions were compared only in the counties of Uusimaa and Lapland, the counties in which the non-response ratio was highest. As considerable deliberation was employed in the census, the classification of the sample may not be quite the same as in the census. The result of the comparison is seen in Tables 13/p. 81 and 14/p. 83 and Fig. 4/p. 84. No significant differences in the occupational distribution were found within Lapland County. But in the county of Uusimaa the ratio of the agricultural and forestry population was significantly too high. This was due to the fact that the sampling method in some communes of Uusimaa County had been incorrectly applied. This was discovered so late that it was impossible to correct the bias. As the non-response ratio in the other counties was lower than in these two the sample in them was assumed to be sufficiently free of bias.

The accumulation phenomenon of birth date entries too may have caused sampling biases. It is possible that this phenomenon also contributed to the over-representation of the oldest age classes and of the agricultural and forestry population.

332. Response bias

As the Main Survey material was collected from the commune of Ranua, which had been subjected to the pilot surveys, a comparison of the results of Pilot Survey II (bookkeeping survey), termed Survey K in the following, and those of the Main Survey (Survey H) was possible. The result of this comparison is shown in Table 15/p. 86 and Fig. 5/p. 89. The 0-hypothesis test with the t-values of STUDENT'S distribution indicated that in the \bar{p}_j estimates of the whole year there was a highly significant difference only in Group 17 (non-working days) and its subgroup 15 (unemployment time). In Groups 12 (total labour input) and 8 (house construction, unpaid work) the difference was probably significant.

The differences in the \bar{p}_j estimates of Surveys H and K emerge more clearly if they are computed as percentages of the estimate of Survey K. Table 16/p. 90 shows these percentual differences for aggregate activity groups 5, 11, 12 and 17 + 19 of Table 15/p. 86. The results indicate that the respondents tended to exaggerate, making the big figures too big and the small figures too small. Although there are exceptions to this phenomenon it is probable that in the interview survey the activity groups that are small, infrequent and occur on scattered days, tend to become too small and the large groups, such as the »input of unpaid labour» too large. In the last-mentioned groups the bias is probably of little importance but the »loss of small groups» must be taken into account in the analysis of the results. It should be emphasized that no very far-reaching general conclusions can be drawn from the comparison surveys of Ranua commune. The relatively good uniformity of the results of Surveys H and K, although the Ranua interviews were more difficult than the average, nevertheless helps to support the serviceability of the interview method.

The comparative samples of Ranua were selected as interpenetrating samples of the same population. The same method was applied in the only comparison of the interview and bookkeeping survey of consumption in Norway (Förbruksundersøkelse 1953) known to the present author. Again the compatibility of the results was considered satisfactory; the difference between them and its causes are not analysed in the survey. It is probable that the potential response bias in an interview concerned with annual consumption is greater than in an interview probing into the annual round of activity.

333. Compilation errors

Several check calculations were made for compilation errors; no such errors were found to be visible in the rounded figures of the tables of results.

334. Random variation

3341. Variance of the number of persons

The influence of random variation in sampling observation and computation on the confidence of the estimates computed from the random sample can be established mathematically by computing the standard deviations (standard errors) of the estimates. A condition is that the variances of each sampling unit are independent (HANSEN et al.

1953, p. 91). As the enumerators of the Main Survey totalled nearly 10 000 and the corrections made in the T-forms were predominantly checked by one person, it may be taken that this condition was met. In the following, simple random sample methods are applied to the computation of estimate variances. As the sample was raised by districts it may be taken to correspond to a stratified random sample. Thus simple random sample formulas may give the estimates somewhat but probably not very much too large variances.

Estimates of two fundamental types were computed from the sample of the Main Survey. In one type the computation unit was the individual (example: Table 17/p. 97), in the other a man-day (example: Table 18/p. 99).

The confidence limits of the former estimates can be determined by qualitative analysis. Good approximate values of the absolute variances of estimates μ' and ψ are obtained from formulas 5, 6, 7 and 8/pp. 95—96 (HANSEN et. al. 1953, pp. 125—126). To make sure, a requirement of $m \leq 100$ and $\mu \geq 35$ was made of the use of formula 8/p. 96 in the results of the Main Survey (cf. HANSEN et. al. 1953, p. 131 and COCHRAN 1953, p. 41). These formulas are based on the assumption that the probability distribution of variables φ and ψ is binomial. Assuming that POISSON'S distribution is involved, we get formulas 6 a, 7 a and 8 a/p. 96 (WEATHERBURN 1947, pp. 47—48). The theory of binomial distribution is more correct, but the last-mentioned formulas provide an approximately correct result if formula 7 a is only employed when $\psi \varphi \geq 0.2$ and formula 8 a only if $\psi \geq 0.2$. When these requirements are not met but $\nu_{m'}^2$ in formula 5/p. 95 is ~ 0 , formula 7/p. 96 assumes an approximate form of 7 b/p. 98. By way of an example, the standard errors of u' estimates computed from formulas 7, 7 a and 7 b are shown in Table 17/p. 97. The result shows that in this case the simplest formula, 7 a, is very serviceable.

3342. The variance of the estimated size of activity group

A second main group of computed estimates consists of those which in the form of either the absolute figures x'_j or the ratios $\bar{p}_j = x'_j / z'$ indicate the distribution into activity groups j of the sum of week-days z' of the population subgroup m' . An example of these results is Table 18/p. 99. So far only the internal variance in stratum m' is discussed.

Although this case involves in principle a qualitative analysis, its formulas cannot be applied as written since the unit (man-day) whose characteristic is in question is not a sampling unit (YATES 1949, p. 195). Hence the variances of estimates x'_j and \bar{p}_j should be computed from formula 1/p. 37.

As a very large number of these estimates are given in the results it was impossible to compute the standard errors for all of them in the way indicated, nor was their publication feasible for technical and cost reasons. So an attempt has been made in the following to develop a method by which the variances of the said estimates can be assessed from the relevant estimates themselves.

The method is derived from the following statement. Assuming that z_i is a constant a and that in all forms $x_{ij} = a$, the question is one of ordinary qualitative analysis.

and thus the internal variance of stratum m' , $v_{x'_j}^2 = \bar{q}_j/m\bar{p}_j$. As in this case the standard deviation of variable x_{ij} is the greatest possible (assuming that \bar{p}_j remains constant), it can be concluded that if the formula is given the general form $v_{x'_j}^2 = c_j^2 \bar{q}_j/m\bar{p}_j$, this formula with the value 1 inserted for coefficient c_j gives variances which the true variances do not exceed even if the requirement that $x_{ij} = a$ is not met. In reality the requirement that $z_i = a$ was not fully met in the Main Survey, but calculations indicated that $v_{z_i}^2$ in all cases was so close to 0 that it may be disregarded. As $c_j^2 = s_{p_{ij}}^2/\bar{p}_j \bar{q}_j$, the coefficient of an activity group of the sample subgroup, c_j , is only dependent on the standard deviation and average mean of variate p_{ij} , in other words on the shape of the percentage frequency distribution curve of variable p_{ij} .

By quantitative analysis coefficient c_j was computed for numerous activity groups of the various subgroups of the sample. Tables 19/p. 102, 20/p. 106 and 21/p. 108 give some of the results of these computations. By the least squares method the empirical adjustment value, c_t , was also computed for several c_j coefficients from formula 11a/p. 104 where r indicates the r th $(\bar{p}_j, s_{\bar{p}_j})$ pair and g the number of the pairs. In Fig. 6/p. 105 are drawn the curves of the standard error approximation equation, $s_{\bar{p}_j} = c_t \sqrt{\bar{p}_j \bar{q}_j/m}$, for some c_t coefficients, and the computed standard errors are plotted. It can be concluded from the graph that c_t coefficients 0.55, 0.75 and 0.90 in these cases give satisfactory approximate values of the standard errors.

The correlation between coefficient c_j , and the percentage frequency distribution of variables p_{ij} , the quality of group m , the size of the district surveyed and type of activity group was also studied both from theoretical distributions and empirically from the material. Some frequency distributions of the material are shown in Table 22/p. 110 and Fig. 7/p. 111. The conclusion reached was that by using c coefficients 0.40, 0.55 and 0.75, observing certain simple rules, satisfactory approximate values of standard errors can be computed for the annual characteristics \bar{p}_j and x'_j . For monthly characteristics, c coefficient 0.9 should be used.

Taking into account also the variance between strata m' , $v_{x'_j}^2$ is obtained from formula 12/p. 114. If $(1 - \varphi)/m$ is so small that it can be omitted and if in addition $\bar{p}_j \gtrsim 20\%$, the shortened formula 12 a/p. 114 may be used. The absolute variances for estimates x'_j and \bar{p}_j are obtained, accordingly, from formulas 13, 13 a, 14 and 14 a/p. 115. When $c^2 \sim (1 - \varphi)$, shortened formulas 12 a and 13 a can be used irrespective of other rules.

Table 18/p. 99 gives the standard errors $s_{x'_j}$ computed from formulas 13 and 13 a/p. 115. The standard errors computed both ways are approximately the same in most of the activity groups, which is due to the fact that in this case $(1 - \varphi) \sim 0.55^2$.

335. Comparison between the results of the Main Survey and results based on other sources of information

Table 23/p. 117 compares the results of the Main Survey for the use of manpower by the male members of farmer families running farms with 0.25—4.99 hectares under plough with the corresponding results of a sample in the agricultural census. As the figures do not refer to an identical labour force group and as the interview method in the census of agriculture was less detailed than that used in the Main Survey, the figures are not fully comparable. It is probable that the labour input figures of the agricultural census include unemployment, holiday, illness and other non-working time, as well as the labour input to hunting, fishing and handicraft. If these points are borne in mind, the figures do not differ significantly, especially as the results of the agricultural census have also been estimated from a sample. Unfortunately the confidence limits of this sample have not been computed.

Table 24/p. 121 compares the Main Survey result regarding the numbers of «professional» paid forest and floating workers with data obtained from the population census.

The figures of the Main Survey refer to the number of men whose input of paid labour to forestry and floating exceeded 80 man-days in 1950. The figures of the population census include only men who themselves gave their main occupation as paid forest workers or their input at a minimum of 4 months of paid labour on forestry and floating. In some cases even general workers with at least a 2 month input into forestry and floating were included in the group. Thus the comparability of the figures is very questionable and the differences do not really prove anything.

In 1950 the Ministry of Communications and Public Works collected from employers by means of a partial survey covering half the country information on the number of forest and floating workers as at the end of each month. In the Ministry the figures were raised to apply to the entire country by doubling them. The present author multiplied the arithmetic mean of the figures for the beginning and the end of each month by 19 (the average monthly labour input of a forest worker according to PÖNTYNNEN 1936), which gave the figures presented in column D of Table 25/p. 123. They were compared with the corresponding labour input data of the Main Survey. The result shows that the figures obtained by the Ministry are regularly smaller than those of the Main Survey, as expected. It is difficult to obtain information from all employers and besides forest owners' own fellings for delivery sales are almost completely omitted in the statistics of the Ministry. With a few exceptions, the ratio between the figures is nevertheless fairly constant and hence the seasonal fluctuation percentages are about the same in both investigations.

The labour input figures for the felling of wood for sale have been compared in the same manner in Table 26/p. 124. The conformity is considerably smaller. As the Ministry's figures must be regarded as rather unreliable, the differences are not proof of bias in the Main Survey.

In Tables 27/p. 126 and 28/p. 127 the Main Survey labour input figures for relief works have been compared with the approximated labour input to relief works of unemployed men in rural communes according to the Ministry of Communications and Public Works. Since the latter are based upon a total material and they must be taken as fairly reliable the comparison shows a bias in this respect in the Main Survey. The

reason might have been partly a response bias (loss of small groups) and partly also the abstruseness of the problem set. It has also been noted that the rural population does not always know if it is working in a relief camp or if the work-site is just an ordinary »free» one. The results of the Main Survey were corrected in this respect to comply with the figures of the Ministry.

34. General assessment of the method employed in the Main Survey

In connection with the population census of 1950 a systematic sample of all male persons, aged 15—64 and living in rural communes on Dec. 31st, 1950 — except for some groups in institutional care — was collected for a survey of rural manpower in 1950. The results of the study and the experience gained in the treatment of the material show that the sample should have taken from all male persons over 13 years in rural communes and, to facilitate raising, preferably also institutional dependents, in order to ensure attainment of the object.

In sampling the birthdays 1 and 15 were used as characteristics of the sampling units to replace random numbers, the respondent's statement on his birth date being accepted in each case. The sampling method is serviceable in principle. A factor to take into account in the choice of suitable birth dates, however, is that listed birthdates have a tendency to accumulate especially on the first day of the month and probably on all dates divisible by 5. This naturally makes it more difficult to raise the estimates of the sample and, besides, may cause biases difficult to control. Provided only birthdays not affected by the accumulation tendency are used and if it is possible to determine the role of non-responses, the raising of such a sample, even without knowing the size of the population, is possible by using a theoretical raising factor calculated from the number of birthdays.

The observation method used was interviews about the annual round of activity in retrospect, employing the week-day as the unit and dividing the results into activity groups by calendar months. Studies of the reliability of the method showed that figures given in the responses may in individual cases differ considerably from the actual round of activity. These variations are mostly random and their effect upon the accuracy of the estimates can be determined by standard error calculations. However, a systematic loss is observed in the small activity groups. This response bias is likely to reduce the activity groups other than labour input in the results, especially unemployment time and possibly the smallest labour input groups. It raises the absolute groups of own labour input in particular but the relative bias, judging by corrections, is negligible in them. As, in spite of the above, the amount of »open» unemployment according to the Main Survey was approximately 4 times the corresponding figure computed from the unemployment card index, this seems to indicate that the effect of response bias cannot reduce the serviceability of the results to any appreciable degree.

Non-response bias in the sample was studied only for age distribution and occupational distribution. The occupational distribution of the sample in Uusimaa County where the proportion of non-response was exceptionally high, seems to favour the agricultural and forestry occupational groups. For this reason caution is called for in the application of the characteristics computed from the Uusimaa County sample.

In the County of Lapland sample no bias was observed. The high degree of non-response arose primarily from the too varied tasks imposed on population census enumerators. This suggests that it would be better to collect the data for labour force surveys separately.

Efforts were made to detect possible biases in the results also by comparing the computed estimates with the corresponding results of other surveys. A bias was found and corrected in the relief labour input results. The differences observed in other comparisons were probably due more to bias in the object of comparison than in the Main Survey.

Analysis of the effect of random variation on the accuracy of the results showed that the standard errors are computable with sufficient accuracy from estimates.

The interview form employed in the survey and the punched card method developed for the treatment of the material met the requirements made satisfactorily. However, on the basis of experience gained corrections can be made in certain details.

Computing n from formula 12/p. 114 gives formula 15/p. 129. For the activity group with a \bar{p}_j estimate of 0.1 we obtain a sample size of $n = 2284$ with a c coefficient of 0.75 and confidence requirement of $v_{x'j} = 0.05$, provided $\varphi = 0.9$. These confidence limits can be considered adequate for practical purposes. The size of the sample required for a certain confidence requirement of the \bar{p}_j ratios can be computed from formula 16/p. 131. With the same conditions as above, we obtain $n = 2240$.

As in the 1950 rural labour force survey the above characteristics were computed for very small subgroups of the sample also and, in addition, for small regions consisting of some few communes, its sample cannot be considered too large for this purpose. For similar future surveys at relatively short intervals to find out the trend data relating to the whole country will suffice in most cases and, provided the use of a simple random sample is possible, a sample of some 2 500—3 000 persons will be adequate. If the frame (registration list) for the selection can be divided into two strata, with the white-collar and factory workers in the one and all the other persons of the frame in the other stratum, the accuracy of the results is likely to increase somewhat. The accuracy of confidence calculations in particular may increase (COCHRAN 1953, pp. 27—30).

35. Cost of the Main Survey

According to the recommendations of the Statistical Office of the United Nations (Preparation ... 1950), which the present paper has tried to follow, a report on a sample survey must include data on the costs also. Apart from printing expenses, the total cost of the Main Survey up to Dec. 31, 1953, amounted to 8 165 364 marks, including material collection expenses totalling 3 449 525 marks (75 marks per form) and the salaries of the permanent staff totalling 1 779 675 marks. The balance was spent on material, renting the machines and paying the salaries of occasional staff. The checking of forms cost mk 5.06 per form, their codification mk 6.42 per form, the punching of person cards including machine rent mk 2.93 per card, punching of time cards by the mark sensing method including machine expenses mk 1.57 per card. Piece-work rates and strict supervision of the standard of the work were applied in the treatment of the material. The experience gained of this method was good.

4. Methods of surveying labour input, unemployment time and structure of the labour force in the light of experience gained

The object of the 1950 rural labour force survey was to obtain a sufficiently reliable picture of three subjects of differing character:

(1) The labour input of the male rural population during the observation year, its distribution by social groups, occupations and districts, and the seasonal fluctuations in the structure of the labour input.

(2) The unemployment time of the male rural population, and also the periods of underemployment (hidden unemployment) and its seasonal variation.

(3) The number of male workers engaged for a shorter or longer period during the observation year in certain rural occupations, especially forestry and agriculture, the composition of these labour force groups and the significance of the principal or subsidiary occupation in the worker's annual round of activity.

The labour input of an industry, e.g. forestry, may be measured from data derived either from enterprises (employers) or workers (employees). The former method is generally considered cheaper, mainly because society may instruct employers to release statistical data of public interest directly to research institutions free of charge. This accounts for the fact that the general tendency has been to base complete enumerations on enterprises.

Statistics collected by mail are in fact considered to supply sufficiently accurate data on labour input to the manufacturing and mechanical industry. Such data can be based on information obtained from the payrolls of the enterprises concerning the number of man-hours worked. In agriculture and forestry and certain related fields of production, however, the case is different. The number of enterprises is very great, their average size small and to a very large extent they employ a labour force comprising the owner family of the enterprise. Records are seldom kept. In addition, it is difficult for the employer to gauge the labour input of a forest labour force working on a piece-rate basis. It is difficult to obtain information by mail. Even the listing of enterprises is laborious.

Although no attempts have been made to date in Finland or, as far as is known, in other countries to obtain information from enterprises on the labour input of the total agriculture or forestry, such information probably can be collected by using sufficiently well planned samples and observation methods. The seasonal fluctuation in labour input may also be estimated by material collected from enterprises.

The material collected from enterprises is less adaptable for an estimate of the total labour input, its seasonal fluctuation and its structure throughout the country or in any particular district. Labour force statistics collected from enterprises hardly ever cover the total labour force. To make the material divisible into subgroups by age, domicile or social group is more laborious still. The total national male labour input, its structure and seasonal fluctuation is thus assessable most simply, cheaply and accurately by means of a sample selected from the population. The distribution of labour input by the residential districts of the labour force is also most reliably estimated by this method. But should a detailed description be required of the distribution of the labour input of some individual industry by districts of utilization, by size classes

of enterprises or by ownership groups, the material collected from enterprises must be resorted to.

The International Labour Office prefers a sample selected from the population to institutional statistics as a method of *surveying unemployment*. The observation method applied in the 1950 survey of the rural labour force also gives some idea of the under-employment time. In measuring the amount of unemployment time instead of unemployed persons it has at least a chance of covering the short periods of unemployment even better than the labour force survey methods employed in the USA and certain other countries.

The total number of persons employed in a particular industry, its distribution on the basis of population characteristics and the duration of the working season of these persons can be estimated only from a sample selected from the population. It is impossible for an enterprise to observe the annual round of activity of workers who have left the enterprise in the course of the year.

If all the study problems outlined here are to be solved simultaneously the best source of information is a sample taken from the population. This sort of sample must be taken for the third of our subjects of study and it supplies at the same time at but little additional cost the data required for the other two subjects. The object for measurement is then the round of activity of the persons in the sample. Among the observation methods, that of interview as to the annual round of activity was found to be both sufficiently accurate and the cheapest. Its greatest drawback is that labour input data for the early part of the year are not obtained until the following year. To eliminate this drawback the observation period may be split up, say into three-month periods. However, the expenses increase almost in proportion to the rise in the number of interviews. If the survey is expected to supply information on the total amount and structure of the labour force engaged during the year e.g. in forestry the observation of the total annual round of activity must be based on identical sampling units.

A national, though not regional, survey of the male labour force like the 1950 survey of the rural labour force in Finland may, as long as no general-purpose sample exists, be carried out on the basis of such a small simple random sample that it might be possible to make it an annual, even a quarterly event. An inquiry into the round of activity during the past quarter might be made every three months, from a different sample each time, and an inquiry into the annual round of activity of the past calendar year from a sample selected after the close of the year. This would give us quarterly information on the labour input and unemployment within each branch of economic life, and after the end of the calendar year it would also be possible to estimate from the most recent material the total amount and structure of the labour force engaged in the most important branches of economic life.

LIITTEET – APPENDICES

Henkilökohtainen tiedustelu

Kaikki 17 - 65 v. miehet.

Pes. Purtille

16
2
2
D

4. Työajan käyttö 1/7 1948-30/6 1949.

Työkausijako:

I	1/7 - 30/9 1948	kesä	92	pv.
II	1/10 - 15/11 1948	syksy	46	"
III	16/11 - 31/12 1948	alkutalvi	46	"
IV	1/1 - 31/3 1949	keskitalvi	91	"
V	1/4 - 15/5 1949	kevät	45	"
VI	16/5 - 30/6 1949	kesä	46	"

	I Kesä 1948	II Syksy 1948	III Alkutalvi 1948	IV Keskitalvi 1949	V Kevät 1949	VI Kesä 1949	Yhteensä pv.
Metsätyö, hakkuu	14	16	91	7			21
" ajo		86	91				127
" muut					12	16	28
Uitto ja lauttaus					10		10
Maataloustyö (muille)							
Tietyö (ei työttömyystyöt)							
Saha- tai lastaustyö							
Rakennus- ja puusepän työ						16	16
Muu ansiotyö (ei työttömyystyöt)							
Tulkiset tehtävät, koulu, asepalv.							
Yht.pv.viljem.tai tontin ulkop.		14	36	91	29	32	202
Oma metsätalous, kotitarvepuu, hakk. ja ruij.	14	16					30
" " myyntihankinnat							
" " pystyynmyynnit							
" " muut työt							
Rakennustyö omalla kiinteistöllä							
Uudisviljely, ojitus ym.	78						78
Kotiteollisuus myyntiä varten							
Oma poronhoito							
Metsästys ja kalastus							
Oman maatalouden työt		16			16	14	46
Yht.pv.viljelmällä tai tontilla	92	32			16	14	154
Työttömyystöissä							
Varsinainen työttömyys							
Puolittain työtön aika (redusoitu)			10				10
Sairaus							
Vapaapäivät (ei pyhä- ja sunn.pv.)							
Yhteensä			10				10
Yhteensä pv. työkaudessa	92	46	46	91	45	46	366

Täytetty I esitutkimuksen ajankäyttölomake

Completed account of annual round of activity, Pilot Survey I

Mihin käyttite vuoden 1950?		Verokirjaa ym. apuna käyttäen on jaettava kuukausien arkipäivien summa (rivi 29) alla olevien ajankäyttölajien kesken (malli viimeisellä sivulla)	Rivi
I Muuhun kuin työhön käytetty aika	Koulussa, opiskelemissa tai asevelvollisena		01
	Vapaaehtoiset loma- tai vapaapäivät	Myös kyläily- ja virkistys- tai huvimatkat, vain arkipäivät	02
	Sairaana	Tähän tulee myös sairaan omaisen hoitamiseen ja matkoihin lääkärin luo käytetyt päivät	03
	Hukkapäivät	Työnhakumatkoihin, työpaikan ja kodin välillä tehtyihin matkoihin tai epäedullisen sään takia hukkaan menneet kokonaiset arkipäivät	04
	Työttömänä ansiotyön puutteen tähden	Työttömyystyöt viedään ao. työläjien kohdalle ja myös riville 31 Lakkopäivät riville 06	05
	Työttömänä lakon tähden	Tähän tulevat vain sellaiset lakkopäivät, jolloin asianomainen ei tehnyt työtä Lakon aikana tehdyt työt merkitään ao. riveille	06
	Muu aika, jota ei ole käytetty työhön. Selitys riville		07
II Vieraan palveluksessa tehdyt työt	Metsätyössä, kotitarvepuuta hankkimassa	Vieraan maatilalla tekemässä tai ajamassa poltto-, rakennus- tai muuta kotitarvepuuta tämän maatilalla omaa tarvetta varten	08
	Metsätyössä, muun kuin kotitarvepuun tekotyössä vieraan metsässä		09
	Metsätyössä, muun kuin kotitarvepuun ajotyössä vieraan metsässä		10
	Metsätyössä, muussa kuin teko- tai ajotyössä	Esim. päivätöissä kuten merkitsemässä, leimaamassa, metsänhoitotöissä, metsäojien kaivussa tms.	11
	Uitto-, erottelu- tai lauttaustyössä	Tähän tulevat myös uittolaitteiden rakennustyöt (ks. myös rivi 18)	12
	Pyöreää puutavaraa lastaamassa tai autolla kuljettamassa		13
	Uudisraivaustyössä vieraan maatilalla	Peltojen tai lautumien raivaus	14
	Asuin- tai talousrakennuksen uudisrakennustyössä vieraalle	Tähän luetaan uusien rakennusten teko, vanhojen uudelleen rakentaminen tai laajentaminen sekä sen yhteydessä tapahtuvat työt kuten muuraukset tai tarveaineiden sahaus rakennuspaikalla tai muualla lähistöllä olevalla sahalla	15
	Rakentamassa julkista rakennusta	Esim. koulua, sairaalaa, tehdasta tai voimalaitosta	16
	Vieraan maatilatalouden töissä	Sellaisissa töissä vieraan maatilalla, jotka eivät ole metsä-, uudisraivaus- tai uudisrakennustöitä	17
	Teiden, rautateiden rakennustyössä tai uittoväylien perkaustyössä		18
	Muissa kuin edellä mainituissa töissä vieraan palveluksessa. Myös henkiset työt	Virhettä ei tule, vaikka kaikki henkiset työt (isännän omalla maatilallaan suorittamaa johtotyötä lukuunottamatta) merkitään tälle riville	19
	III Omat työt	Käsityössä tai pienteollisuudessa	Oma työ (siis ei vieraan palveluksessa tehty) vieraan tarvetta varten Sepän, suutarin yms. työt, työ omassa myllyssä, omalla sahalla yms. Työnjohto riville 28
Metsästämissä tai kalastamassa tai porojen hoidossa			21
Kotitarvepuun teossa tai ajossa		Omaa tai oman tilan tarvetta varten. Tähän luetaan poltto-, rakennus- ja muiden kotitarvepuuiden teko ja ajo	22
Omassa metsässä myyntipuun teossa tai ajossa		Sekä hankinta- että pystyynmyynnit. Vaikka työskentely omassa metsässä olisi tapahtunut puiden ostajan palkkaamana, tulee se tähän	23
Omassa metsässä muissa metsätöissä		Ne työt omassa metsässä, jotka eivät sovi edellä kohtiin 22 ja 23	24
Omissa uudisrakennustöissä		Selitys rivillä 15	25
Omissa uudisraivaustyössä		Peltojen tai lautumien raivaus	26
Muissa omissa ruumiillisissa töissä		Omissa ruumiillisissa töissä, jotka eivät sovi edellä kohtiin 20-26 Tähän luetaan omien viljelysten ja kotieläinten hoito (ei kuitenkaan porojen), omien rakennusten, kaluston ja muiden varusteiden kunnossapito, omien maataloustuotteiden myynti ja myyntipaikalle kuljetus, kotiloukutyöt yms.	27
Omat johto-, toimistoym. henkiset työt		Viljelmän isännän johtotyöt luetaan kuitenkin kohtiin 22-27	28
		Yhteensä päiviä	29
A Luppoaika Omilla töillä peitetty työttömyys	Tämä koskee vain niitä, joista merkittiin päiviä yllä III ryhmään (rivit 20-28)	Kuinka monta riveille 20-28 merkityistä päivistä olisi asianomainen halunnut ja omien töiden siitä kärsimättä - tai joutumatta toisten suorittaviksi - voinut käyttää ruumiilliseen ansiotyöhön, esim. metsätyöhön, jos sitä olisi jatkuvasti ollut saatavissa niin lähellä, että hän olisi voinut siitä suorittaa kotoa käsin Merkittävät päivät myös tälle riville kuukausittain	30
B Työttömyystyöt	Tämä koskee vain niitä, joista merkittiin päiviä yllä II ryhmään (rivit 08-19)	Oliko joku osa ylle merkityistä ansiotyöistä virallisen työttömyyskortiston välityksellä saatavia työttömyystöitä? Jos oli, merkittävät työttömyystöihin käytettyjen päivien lukumäärä myös tälle riville kuukausittain	

Vuoden 1950 kuukaudet												Laskija ei täytä	
I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII		
												01	
			4							1	2	7	
												03	
3	2	4		5					3	3	2	04 22	45 110 11
												05	09 - 19
												06	44 12b
												07	15+16 -
												08	13
22	22	12							20		16	09 92	→ 10
												10	→ 14
												11	-
												12 18	→ 02
				18								13	43 16
												14	272 28
												15	34+36+38
												16	42 17
												17	153 16
												18	33+36
												19	41 18
												20	162 17
												21	34+40
												22	40 19
												23	38+08 -
												24	39 20
												25	119 12
			19							22	2	22 43	32+38
												26	38 21
												27	14+17 -
												28	37 22
												29	18 02
												30	05+30+31
						22	19	25				31	36 23
												32	110 11
												33	35+12 24
												34	35 24
												35	92 24
												36	110 11
												37	35+12 24
												38	35 24
												39	92 24
												40	110 11
												41	35+12 24
												42	35 24
												43	92 24
												44	110 11
												45	35+12 24
												46	35 24
												47	92 24
												48	110 11
												49	35+12 24
												50	35 24
												51	92 24
												52	110 11
												53	35+12 24
												54	35 24
												55	92 24
												56	110 11
												57	35+12 24
												58	35 24
												59	92 24
												60	110 11
												61	35+12 24
												62	35 24
												63	92 24
												64	110 11
												65	35+12 24
												66	35 24
												67	92 24
												68	110 11
												69	35+12 24
												70	35 24
												71	92 24
												72	110 11
												73	35+12 24
												74	35 24
												75	92 24
												76	110 11
												77	35+12 24
												78	35 24
												79	92 24
												80	110 11
												81	35+12 24
												82	35 24
												83	92 24
												84	110 11
												85	35+12 24
												86	35 24
												87	92 24
												88	110 11
												89	35+12 24
												90	35 24
												91	92 24
												92	110 11
												93	35+12 24
												94	35 24
												95	92 24
												96	110 11
												97	35+12 24
												98	35 24
												99	92 24
												100	110 11

Täytetty päätukimuksen lomake (T-lomake), sivu 3

Completed form of the Main Survey (T-form), page 3

Päätukimuksen lomake (T-lomake), sivu 2

Form of the Main Survey (T-form), page 2

English translation of the main headings of the T-form

Page 1

(Figures refer to the numbers on left hand side of the page.)

1. Census commune.
2. In which commune were you living on Dec. 31, 1950.
3. Name.
4. Address.
5. Date of birth. Day — month — year.
6. Main occupation in 1950.
7. How many hectares of land under plough, garden and forest are there on the farm you last cultivated in 1950 for your own account? If you did not run a farm in 1950 for your own account, how many hectares of land under plough, garden and forest were there on the farm or place where you lived Dec. 31, 1950? Put X or state the number of hectares in one square in each of the three tables below. The lowest square also covers a case with no land at all.
8. Do you own or rent the farm or site or are you a resident member of the family which owns or rents the farm or site?
9. Were you employed by some other person in 1950?
Are you usually employed by one or more employers during the year?
10. Was the amount of work you did in 1950, as a result of illness or disablement, less than half of what you would have done if healthy?

Page 2

Annual round of activity table

(Numbers of activity groups refer to the mid-column of the page.)

State with the aid of your tax book or something similar how you were occupied on each week-day of the month. See example on the last page.

I Non-working weekdays.

- 01 School studies and military service.
- 02 Holidays or days off.
- 03 Illness.
- 04 Non-working days — whole week-days wasted on seeking employment, trips between work-site and home, or because of inclement weather.

- 05 Unemployed through lack of work.
- 06 Unemployed through a strike.
- 07 Unemployed for other reasons.

II Work when employed by another person.

- 08 Felling or hauling wood for domestic purposes.
- 09 Forest work, felling wood for other than domestic purposes in someone else's forest.
- 10 Forest work, hauling wood for other than domestic purposes in someone else's forest.
- 11 Forest work other than felling or hauling.
- 12 Floating.
- 13 Loading or motor transport of roundwood.
- 14 Land clearance for agriculture.
- 15 Work on buildings other than public buildings.
- 16 Work on public buildings.
- 17 Agricultural work.
- 18 Work on roads, railways and floatways.
- 19 Work not mentioned above. All white-collar work.

III Work for own account and unpaid family work.

- 20 Handicraft or small industry.
- 21 Hunting and fishing or reindeer breeding.
- 22 Felling or hauling wood for domestic use.
- 23 Felling or hauling wood for sale in own forest.
- 24 Other forest work in own forest.
- 25 House construction.
- 26 Land clearance for agriculture.
- 27 Other manual work than the above.
- 28 White-collar work for own account.

29 Total.

A. Underemployment (unemployment hidden by self-employment).

- 30 Applies only to Group III above.
How many of the days entered against 20—28 would the respondent have *liked* and, without his own work suffering or falling on others, *been able* to spend on *paid* manual work, e.g. forest work, had it been continuously available close enough for him to live at home while thus employed. Please enter the number of days here *too* by months.

B. Relief work.

- 31 Applies only to Group II above.
Was any of the paid work entered above relief work obtained through the official relief work card index? If so please enter the numbers of days spent on relief work on this line *too* by months.

T-lomakkeiden esitarkastajille annetut tärkeimmät ohjeet

Instructions to the workers checking the T-forms

I. TARKOITUS

Esitarkastuksessa merkitään lomakkeen etusivun koodit kohtiin 7 (Pinta-alat) — 10 (Työkyky). Samalla tarkastetaan, että lomake on oikein täytetty vertaamalla etusivun ja sisäaukeaman tietoja toisiinsa. Puutteellisesti täytetyt lomakkeet, joita ei voida siinä esitettyjen tietojen perusteella ohjeiden mukaan täydentää, tai puutteellisina hyväksyä, erotetaan pääkäsittelystä ja siirretään jälkitarkistukseen, (etusivulle T-kirjaimen oikealle puolelle tehdään merkki 0). Jälkitarkistuksessa ne vielä uudelleen tarkastetaan ja tarpeen vaatiessa palautetaan kuntiin täydennystä varten. Lomakkeet, joissa on ristiriitaisia tai epätodennäköisiltä näyttäviä tietoja, siirretään myös jälkitarkistukseen.

II. ESITARKISTUKSEN SUORITUS OIKEIN TÄYTETYSSÄ LOMAKKEESSA

A. Etusivu

Kohta 7: pinta-alatiedot. Jos pellon pinta-alaa tarkoittavassa ruutusarjassa rasti on ruudussa 3 = 5—9,9 ha, merkitään oikealle koodiruutuun 6 (= Pelto) koodi 3. Jos samassa ruutusarjassa on ruudussa 1 = 25 ha tai enemmän esim. luku 35, merkitään vastaavaan koodiruutuun koodi 1. Samalla tavalla merkitään puutarhan ja metsän pinta-alojen koodit.

Kohta 8: Oletteko tämän viljelmän... Verrataan vastausta merkintään ammatista kohdassa 6 sekä myös sisäisivun päivämerkintöihin osastossa III »Omat työt». Jos ristiriitaa ei ole esim. siten, että vastaus olisi »En ole», mutta omia töitä maataloudessa silti on huomattava määrä (ks. jällempänä »Täyttämismvirheet»), merkitään asianomaiseen koodiruutuun (9) rastiruudun ilmaisema koodi (esim. vastaus »Olen», koodi = 1).

Kohta 9. Katsotaan onko sisäisivulla päivämerkintöjä osastossa II Työt vieraan palveluksessa. Jos ao. merkintöjä ei ole, tulee tähän kohtaan aina koodi 1 (oikein täytetyssä lomakkeessa silloin on myös rasti ruudussa 1). Oikein täytetyssäkkin lomakkeessa voi tällöin olla lisäksi rasti ruudussa 2 tai 3. Nämä rastit jätetään tässä tapauksessa huomiotta.

Jos sisäisivulla sen sijaan on päivämerkintöjä osastossa II, merkitään kohdan 9 koodiruutuun koodi 2 tai 3 riippuen siitä, kummassa ruudussa rasti on.

Kohta 10. Katsotaan sisäisivulta riviltä 03 = Sairaana, montako sairauspäivää T-miehelle on merkitty. Jos niitä on vähemmän kuin 150 ja tässä kohdassa etusivulla

on rasti ruudussa 2, merkitään kohtaan koodi 2. Jos niitä on 150 tai enemmän (jolloin myös rastin tulee olla ruudussa 1) merkitään koodi 1. Jos rasti on ruudussa 1, vaikka sairauspäiviä on vähemmän kuin 150, tulee harkita, onko kohdan 10 kysymykseen todella vastattu oikein. Yleensä on silloin myös joko etu- tai takasivulla jokin selitys asianomaisen työkyvyttömyydestä, joka vahvistaa merkinnän oikeutuksen. *Aina on tarkoin harkittava*, onko syytä merkitä tähän kohtaan työnsuorituskyvyltään erittäin heikon miehen koodi 1. (Ks. »Täyttämismvirheet».)

B. Sisäisivu

Ensimmäinen tehtävä on tarkastaa vaakasuunnassa tehdyt päivämerkinnät riviltä ylhäältä alas. Samalla on pidettävä mielessä T-miehen ammatti. Jos lomake on oikein täytetty, tulee merkintöjen vaikuttaa etusivun tietoihin verrattuna todennäköisiltä. Jos sisäisivulle on kirjoitettu selityksiä, on todettava, ovatko selitykset yhtäpitäviä päivämerkintöjen kanssa. Erytystä huomiota tulee kiinnittää riveillä 30 ja 31 mahdollisesti esiintyviin merkintöihin.

Toinen tehtävä on pystysuorasti tarkastaa kuukausisarakeittain päivien yhteenlasku. Summarivin (29) yläpuolella olevien lukujen summan tulee sarakeittain olla = riville 29 painettu luku (tai sille kirjoitettu luku, jos pyhäpäivinä on tehty töitä). Riville 30 merkitty luku ei saa olla suurempi kuin vastaavaan sarakeeseen osastoon III merkittyjen päivien summa. Riville 31 merkitty luku ei saa olla suurempi kuin osastoon II merkittyjen saman kuukauden päivien summa.

III. TÄYTTÄMISVIRHEET JA NIIDEN KÄSITTELY

A. Puutteelliset merkinnät

Kohta 7. Jos ainoastaan puutarhan pinta-alatiedot puuttuvat, merkitään asianomaisen koodin kohdalle merkki —. Puute ei siis aiheuta enempää toimenpiteitä.

Milloin T-mies ei ole viljelijä, muitakaan pinta-alatietoja ei tarvita. T-mies ei ole viljelijä, jos lomake täyttää toisen tai molemmat seuraavista ehdoista:

- 1) Jos kohdassa 8 on koodi 2, eikä tämä ole ristiriidassa ajankäyttötaulukon kanssa.
- 2) Jos osastossa III ei ole merkitty riveille 22—27 päiviä enempää kuin 15 yhteensä.

Kohta 8. Ei viljelijältä (ks. edellä määritelmää) ei vaadita kohdan 8 tietoja, jos hän ei ole tehnyt metsä- tai uittotöitä (päiviä riveillä 08—12).

Kohta 9. Jos T-mies on ollut vieraan palveluksessa (päivämerkintöjen mukaan), merkintöjen puute tässä kohdassa on korvattava harkinnalla. Jos päättelemisen on mahdotonta, on koodi = 2.

Kohta 10. Merkinnän puuttuminen tulkitaan merkitsemällä koodi 2, mikäli jokin syy ei nimenomaan anna aihetta merkitä koodia 1.

Milloin jotakin tietoa ei ole annettu ja niiden puuttuminen hyväksytään, merkitään ao. koodiruutuun —. Asianomaisen tiedon arvaaminen ei saa tulla kysymykseen. Sen sijaan sallitaan esim. kohdan 8 tiedon päättelemisen ammatin perusteella (maan viljelijä).

B. Virheelliset ja ristiriitaiset merkinnät

Kohta 7. Kun pinta-alatieto on merkitty numeroin, voi tämä numero sijaita väärässä rastiruudussa. Kodifointi suoritetaan pinta-alatiedon, eikä numeron sijoituspaikan mukaan.

Kohta 8. Jos vastaus tässä kohdassa on »En ole», voi sattua, että pääammatti ja vastaus sisäisivulla osastossa III ovat ristiriitaiset. Jos ammatti on »Maanviljelijä» tai »Maanviljelijän poika» ja pääosa työajasta on merkitty »Omiin töihin», voidaan kohdassa 8 merkitä koodi 1 laskijan ilmeisestä virhemerkinnästä huolimatta. On kuitenkin huomattava, että myös esim. vuokralla asuvalta maataloustyöläiseltä tai muulta maatomalta T-mieheltä voi kulua jonkin verran aikaa »omiin töihin» esim. vaimon auttamiseen kotitöissä, omien polttopuiden hankintaan tms. On mahdoton määritellä näiden omien töiden suurinta sallittua määrää. Harkintaa on tässä käytettävä.

Ammattinimitys »maatyöntekijä» tai vastaava ei riitä tekemään myöntävää vastausta kohdassa 8 vääräksi. Monin paikoin nimitetään maanviljelijän poikia maatyömiehiksi. Sitä paitsi hyvin useilla maataloustyöläisillä on omassa hallinnassaan pieni maatila tai tontti. Kohdassa 8 esiintyvät virhemerkinnät ja ristiriitaisuudet saa korjata vain, jos virheen laatu varmasti voidaan todeta. Muussa tapauksessa lomake on siirrettävä jälkitarkistukseen.

Kohta 9. Virhemerkintöjen tulkinta on pääasiassa käsitelty edellä. Jos tässä kohdassa on rasti ruudussa 1 ja joitakin harvoja päiviä on merkitty osastoon II, tämän ei aina tarvitse olla virhemerkintä. T-mies on voinut käydä auttamassa naapureitaan tai merkinnyt riville 19 jonkin käytännöllisesti katsoen palkattoman luottamustoimen kokouspäivät. Tällaisissa tapauksissa merkintä »En ollut vieraan palveluksessa» on katsottava oikeaksi.

Kohta 10. Muutamat heikot laskijat ovat käsittäneet tämän kysymyksen väärin. Rastiin ruudussa 1 on aina suhtauduttava erittäin varoen. Pelkkä korkea ikä ei oikeuta merkitsemään T-miestä työkyvyttömäksi. Jos on perusteltu syy olettaa, että rasti on merkitty väärään ruutuun, on lomake siirrettävä jälkitarkistukseen.

Työpäivämerkinnät. Selvästi väärälle riville tehdyt työpäivämerkinnät siirretään oikealle rivilleen. Jos ajankäyttömerkinnät eivät tunnu vastaavan etusivun (ammattin) merkintöjä, siirretään lomake jälkitarkistukseen. Riville 27 ei tulisi olla merkitty muita omia töitä kuin maatalouden ja kotitalouden juoksevat työt. Mikäli esim. autonkuljettajasta on merkitty hänen tämän ammatin harjoittamiseen käyttämä aika riville 27, siirretään lomake jälkitarkistukseen. Milloin *luppoaikaa* (rivi 30) on merkitty, on tarkoin todettava, että se on tapahtunut lomakkeen ohjeiden mukaisesti. Virheellisiä merkintöjä ei korjata, lomake siirretään jälkitarkistukseen. Sama koskee *työttömyystyö-*merkintöjä rivillä 31.

Jos *päivien riville 29 laskettu summa* on suurempi kuin tälle riville painettu luku, ei osastossa I saa olla ko. kuukauden kohdalla merkintöjä. Mikäli I osastossa tällöin on merkitty päiviä, on näistä päivistä poistettava niin monta, että summaksi tulee painettu luku tai luku, joka on = osastoihin II ja III merkittyjen päivien summa. Päivien poistaminen tapahtuu seuraavassa järjestyksessä:

Hukkapäivät rivi	04
Työttömyys »	05
»	06
Muu aika »	07

Vapaapäivät rivi	02
Sairaus »	03
Opiskelu ym. »	01

Päivien summa ei saa olla suurempi kuin ko. kuukauden kaikkien päivien summa.

IV. ERIKOISTAPAUKSET

A. Hylättävät:

Hylkäysmerkinnällä XXX varustetaan lomake, jos se on täytetty miehestä, joka on joko

a) ollut koko vuoden 1950 puolustusvoimien tai rajavartioston *vakinaisessa* palveluksessa (asevelvolliset eivät siis kuulu tähän) tai

b) on ollut koko vuoden 1950 *hoidettavana* jossakin laitoksessa kuten kunnalliskodissa, sairaalassa tai vankilassa.

Hylätyt lomakkeet erotetaan aineistosta.

B. Ulkomailla työskennelleet.

Koko vuoden tai osan vuodesta 1950 ulkomailla oleskelleiden T-miesten lomakkeet siirretään jälkitarkistukseen.

C. Työkyvyttömät.

Jos lomakkeeseen on perustellusti merkitty kohtaan 10 koodi 1, ei lomakkeessa välttämättä tarvita tämän lisäksi muita merkintöjä kuin ikä sekä merkinnät työhön mahdollisesti käytetyistä päivistä. Kohtiin, joista tiedot puuttuvat, merkitään etusivun koodiruuduissa merkki —.

V. YLEISIÄ OHJEITA

Laskijan merkintöihin saa tehdä muutoksia vain, jos varmuudella voidaan päätellä, että tiedot muutoksen jälkeen ovat oikeat. Epävarmoissa tapauksissa tulee käsittelijän aina kääntyä työn valvojan puoleen tai, mikäli se ei ole mahdollista, siirtää lomake jälkitarkistukseen.

Käsittelijän tulee aina muistaa, että jokainen lomake edustaa 15 samanlaista tapausta.

Myös jälkitarkistukseen siirrettävät lomakkeet on tarkastettava loppuun saakka. Epäselvät kohdat on selvästi kysymysmerkein merkittävä (ks. mallia).

Laskijan tekemät epäselvät numerot on esitarkistuksen suorittajan selvennettävä. Luppoaikapäivät on, mikäli ne ovat rivin 30 yläreunassa, siirrettävä rivin alareunaan, jotta kortistoinnissa ei syntyisi epäselvyyttä siitä, tarkoittavatko numerot muutoksia päivien summaan vai luppoaikaa.

Kun riville 07: »Muu aika, jota ei ole käytetty työhön», on merkitty päiviä ja niiden laatu selitetty, on huomattava seuraavaa: Jos selityksen avulla päivät voidaan siirtää jollekin muulle riville, on se aina tehtävä. Luottamustoimien kokouspäivät siirretään

riville 19. Maataloustyöhön, kalastukseen tms. liittyviin myynti-, osto- yms. matkoihin käytetyt päivät viedään ao. työajain kohdalle.

On huomattava, että työnjohto tulee merkitä riville 19 (ks. lomaketta ao. kohdassa).

Esitarkistuksen merkinnöissä ei saa käyttää mustaa, sinistä eikä punaista väriä, koska laskijat ja laskentatoimikunnat käyttävät näitä värejä. Kaikki merkinnät on tehtävä selvin numeroin.

* * *

Puutteellisten maatilain pinta-alatietojen hyväksyminen eräissä tapauksissa johtui siitä, että pellon, puutarhan ja metsälön pinta-alatiedot hyvin lukuisissa tapauksissa oli jätetty täyttämättä. Useimmiten syynä lienee ollut, että tapauksissa, joissa peltoa, puutarhaa tai metsää ei ollut ollenkaan, asianomainen rasti oli jätetty merkitsemättä. Esim. pinta-alaluokkaan 0—0.24 ei ymmärretty kuuluvan tapauksen, jossa peltoa ei ollut ensinkään. Tätä osittain liian vaikeatajuisesta lomakkeesta johtuvaa puutetta oli mahdoton korjata lähettämällä kaikki asianomaiset lomakkeet muutenkin yli-kuormitetuille laskentatoimikunnille, joiden usein oli vaikea hankkia pinta-alatietoja. Pinta-alatietojen ehdoton vaatimus rajoitettiin sen tähden vain viljelijöihin. Lisäksi päätettiin vähämerkityksiset puutarhan pinta-alatiedot niiden puuttuessa merkitä tuntematon yksi, jollei lomakkeita muista syistä palautettu kuntiin ja laskentatoimikunta pystynyt niitä siellä täyttämään.

Lauri Heikinheimo: Metsätyövoiman tutkimusmenetelmä
(Acta Forestalia Fennica 63.1)

Painatuksen jälkeen havaitut virheet — *Errata*

Sivu 20, 14. rivi ylhäältä: ss. 292—244 — tulee olla — ss. 292—294

Sivu 36, 7. rivi ylhäältä: Jos $z_i = a$, on \bar{p} myös — t.o. — Jos $z_i = a$, on \bar{p}_j myös

Sivu 36, 5. rivi alhaalta: s_x^2 t.o. s_x^2

Sivu 58, 1. rivi ylhäältä: 31. 12. 1949 — t.o. — 31. 12. 1950

Sivu 88, 10. rivi alhaalta: STUDENTIN jakautuman — rivin tulee alkaa tekstin reunaan, koska edellä aloitettu lause tässä jatkuu

Sivu 106, taulukon alla: 3. rivi ylhäältä: sekä II — t.o. — sekä I

Sivu 107, 1. rivi ylhäältä: II esitutkimuksen — t.o. — I esitutkimuksen

Sivu 109, asetelma: merkinnät \bar{p}_j ja c_j ovat 1 cm:n verran liiksi oikealla

Sivu 115, 3. tekstirivi ylhäältä: Estimaatin \bar{p}_j -varianssi — t.o. — Estimaatin \bar{p}_j varianssi

Sivu 130, ylemmän kaavan ja sitä seuraavan sivun osan tulee olla seuraava:

$$v_{n'}^2 = \frac{1-f}{n'f}$$

jossa f on teoreettinen otantaosuus n/n' . Ratkaisemalla n saadaan $n = n'/(n'v_{n'}^2 + 1)$. Kun Suomen maaseudulla n' on suuruusluokkaa $8 \cdot 10^5$, tulee varmuusvaatimuksella $v_{n'} = \pm 0.05$ näytteen suuruudeksi $n \sim 400$ koehenkilöä, joten mainittu varmuusvaatimus likimäärin tulisi täytetyksi koko maaseudun arvioluvuissa $n = 2\,700 - 3\,000$ koehenkilön näytteellä. Laskuesimerkki osoittaa, että tulosten varmuuden kannalta on jokseenkin samantekevää, käytetäänkö todellista vai teoreettista suurennustekijää edellyttäen, että molemmat ovat harhattomat. Jos tyydytään pelkästään suhdelukuihin \bar{p}_j , ei perusjoukon suuruuden tunteminen ole tarpeen. Suhdelukujen \bar{p}_j arviointia varten vaadittavan otoksen suuruus on