

**ETLA**

**ELINKEINOELÄMÄN TUTKIMUSLAITOS**  
THE RESEARCH INSTITUTE OF THE FINNISH ECONOMY  
Lönnrotinkatu 4 B 00120 Helsinki Finland Tel. 358-9-609 900  
Telefax 358-9-601 753 World Wide Web: <http://www.etla.fi/>

## Keskusteluaiheita - Discussion papers

No. 568

Eugen Koev

**PALKKADISKRIMINAATIO**  
**TEOLLISUUDEN TOIMIHENKILÖILLÄ**

**KOEV, Eugen, PALKKADISKRIMINAATIO TEOLLISUUDEN TOIMIHENKILÖILLÄ.** Helsinki: ETLA, Elinkeinoelämän Tutkimuslaitos, The Research Institute of the Finnish Economy, 1996. 73 s. (Keskusteluaiheita, Discussion Papers, ISSN 0781-6847; no. 568).

**TIIVISTELMÄ:** Työssä arvioidaan sukupuoleen perustuvaa palkkadiskriminaatiota (palkkaeroja samasta työstä miesten ja naisten välillä) Suomen teollisuuden toimihenkilöillä. Aineisto on poikkileikkausotos vuodelta 1993 ja paneeli vuosilta 1991 ja 1993. Ne on poimittu Teollisuuden ja Työnantajain Keskusliiton palkkatilastoista. Estimoitujen toimihenkilöiden palkanmuodostusta kuvaavien tilastollisten mallien perusteella konstruointiin palkkadiskriminaatiofunktio.

Palkkadiskriminaatiofunktio on kompakti kuvaus siitä, mistä tekijöistä ja millä tavalla palkkadiskriminaatio riippuu. Saatujen tulosten mukaan palkkadiskriminaatio vaihtelee toimihenkilöiden keskuudessa mm. työtehtävittäin, toimihenkilöryhmittäin ja toimialoittain. Tyyppillinen naistoimihenkilö saa noin 10 % pienempää palkkaa kuin vastaavaa työtä tekevä mies. Palkkadiskriminaation osuus kokonaispalkkaerosta sukupuolten välillä,  $1/3 - 1/4$ , on kuitenkin suhteellisen pieni. Paneeliaineistoon sovellettiin ns. vakioisten vaikutusten mallia (fixed effects model). Menetelmällä haluttiin ottaa huomioon lahjakkuuden ja ahkeruuden kaltaisten ei-havaittujen tekijöiden vaikutuksia palkanmuodostukseen. Vakavien identifioituvuus- ja multikollineaarisuusongelmien ja aineiston ei-kokeellisen luonteen vuoksi työ jäi tältä osin tuloksetta.

**AVAINSANAT:** palkanmuodostus, palkkaerot, sukupuoli, diskriminaatio.

**KOEV, Eugen, WAGE DISCRIMINATION AMONG EMPLOYEES IN MANUFACTURING.** Helsinki: ETLA, Elinkeinoelämän Tutkimuslaitos, The Research Institute of the Finnish Economy, 1996. 73 p. (Keskusteluaiheita, Discussion Papers, ISSN 0781-6847; no. 568).

**ABSTRACT:** The study assesses the male-female wage discrimination (wage difference for the same work) among employees in the Finnish industry. The data used, a cross-section sample from the year 1993 and a panel from the years 1991 and 1993, is from the wage statistics maintained by the Confederation of Finnish Industry and Employers. On the basis of the estimated statistical models, which describe the wage formation of employees, a wage-discrimination function was constructed. The wage discrimination function is a summary of what factors affect discrimination and how. The results obtained suggest that wage discrimination among employees varies, among other things, by work assignment, by employee's category and by branch of industry. A typical female receives about 10 % lower wage than a male doing similar job. The wage discrimination accounts for  $1/4 - 1/3$  of the total wage gap between genders. The so-called fixed effects model was applied when analysing the panel data. The purpose of using this method was to account for the effect of unobserved factors, such as ability and diligence, on wage formation. This part of the study remained without results because of serious identification and multicollinearity problems combined with the non-experimental nature of the data.

**KEY WORDS:** wage formation, wage differences, gender, discrimination

## SAATTEEKSI

Tämä työ on esitetty ja hyväksytty pro gradu-tutkielmana Helsingin yliopistossa huhtikuussa 1996. Tutkielmani ohjaajana ja tarkastajana toimi professori Yrjö Vartia, jota haluan kiittää opastuksesta kärsivällisyydestä ja henkisestä tuesta, joka on ollut tarpeen. Lämpimästi kiitän tutkielman toista tarkastajaa, professori Erkki Koskelaa, positiivisesta palautteesta. Monia työn ongelmakohtia olen pohtinut yhdessä Antti Suoperän kanssa.

Kiitän ETLA:a tutkimusaineistosta, mahdollisuudesta käyttää laitoksen resursseja ja saamastani taloudellisesta tuesta. Erityiset kiitokset VTT Pentti Vartialle, joka kiireistään huolimatta varasi aikaa tutkielmaani varten ja jonka kommentit ja ehdotukset ovat olleet hyödyksi. Tutustuessani tutkimusaineistoon sain paljon apua Juha Kettuselta, jolle myös esitän kiitokseni.

Tehdessäni tätä tutkimusta olen oppinut paljon, mutta paljon on vielä opittavaa. Jo nyt tulee mieleen sanoa, että jos pitäisi aloittaa alusta, tekisin jotkut asiat toisin. Työn puutteista ja mahdollisesti esiintyvistä virheistä tai virheellisistä päätelmistä vastan tietenkin itse.

Helsingissä elokuun 28. päivänä 1996  
Eugen Koev

# SISÄLLYS

1 TAVOITTEET, RAJAUKSET JA SISÄLTÖ	1
1.1 Tavoitteet ja rajaukset	1
1.2 Alan tutkimus Suomessa ja tämän työn yhteydet siihen	5
1.3 Työn sisältö	6
2 TUTKIMUSMENETELMÄ	8
2.1 Lähtökohdat	8
2.2 Kovarianssianalyysin esittely palkkadiskriminaatiotutkimuksen yhteydessä	9
2.3 Palkkamallien logaritmisesta muodosta	14
2.4 Palkkadiskriminaation arviointi	16
3 TUTKIMUSAINEISTO JA MUUTTUJAT	21
3.1 Tutkimusaineisto	21
3.2 Muuttujat	23
3.3 Yhteenveto	35
4 PALKKADISKRIMINAATIOFUNKTION ESTIMOINTI	36
5 PALKKADISKRIMINAATION ARVIOINTI	45
5.1 Naisten ja miesten välisen palkkaeron dekompositio	45
5.2 Palkkadiskriminaatio toimihenkilöryhmittäin ja toimialoittain	47
5.3 Palkkadiskriminaation jakauma tutkimusaineistossa	50
5.4 Yhteenveto	52

6 ERÄIDEN EI-HAVAITTAVIEN TEKIJÖIDEN MERKITYKSESTÄ PALKKADISKRIMINAATION ARVIOINNIN KANNALTA JA NIIDEN MALLINTAMINEN PANEELIAINEISTOSSA	53
6.1 Ei-havaittavat tekijät ja kyvykkyysarhaongelma palkkatutkimuksessa	53
6.2 Paneelaineisto ja vakioisten vaikutusten malli	56
6.3 Vakioisten vaikutusten mallin soveltuvuusongelmat	58
6.4 Empiiriset tulokset	62
7 LOPUKSI	68
LÄHTEET	71
LIITE 1. Palkkadiskriminaatioanalyysissä käytetyt muuttujat ja niiden tunnusluvut v. 1993 poikkileikkausaineistossa	
LIITE 2. Muuttujien ULI, kontvaat ja teknvaat muodostaminen	
LIITE 3. Malli PM1, jossa selittäjänä on työkokemus, ja siihen perustuva palkkadiskriminaatioanalyysi	
LIITE 4. Mallin OLS1 estimointitulokset	

## YHTEENVETO

Naisiin kohdistuva syrjintä työmarkkinoilla on viime aikoina ollut aktiivisen tutkimuksen kohteena. Aihepiiriin liittyy useita erillisiä kysymyksiä, jotka kuitenkin suhteellisen helposti sekoittuvat keskenään. Julkisessa keskustelussa kiinnitetään huomiota siihen, että naisvaltaisten ja miesvaltaisten ammattien välillä on suuri palkkaero. Toinen kysymys on, etenevätkö naiset urallaan samalla tavalla kuin miehet. Kolmas kysymys, jota on selvitetty tässä, liittyy miesten ja naisten palkkaeroihin samantyyppisessä työssä. Tällaisia palkkaeroja on jatkossa kutsuttu palkkadiskriminaatioksi.

Osa sukupuolten välillä esiintyvistä palkkaeroista selittyy sillä, että naiset ja miehet eroavat toisistaan työtehtävänsä ominaisuuksien ja tiettyjen työsuorituksen kannalta keskeisten henkilökohtaisten ominaisuuksien, esim. koulutustason, suhteen. Tällaisista tekijöistä johtuvaa palkkaeroa ei voi tulkita palkkadiskriminaatioksi. Vaikka naiset ovatkin keskittyneet matalapalkkaisiin tehtäviin, ei tämä tarkoita, että he olisivat työtehtävän sisällä huomattavasti huonommin palkattuja kuin mieskollegansa. Toki naisten systemaattinen sijoittuminen matalapalkkaisiin tehtäviin mahdollisesti heijastaa naisiin kohdistuvaa syrjintää, mutta tämä ei ole varsinaista palkkasyrjintää. Kyseessä voi olla naisten pääsyn estäminen (tai vaikeuttaminen) paremmin palkattuihin, perinteisesti miehille kuuluviin tehtäviin. Kuitenkin ammattien jakautuminen nais- ja miesvaltaisiin on monimutkaisen prosessin tulos, eivätkä selvityksessä käytetty aineisto ja menetelmä anna mahdollisuutta tutkia, missä määrin työtehtävien segregointi sukupuolen mukaan on naisten kohtaaman työmarkkinasyrjinnän tulos. Toisaalta, jos naisten pääsyä parempiin palkattuihin tehtäviin estetään, tasa-arvoa edistäviä toimenpiteitä tulisi kohdistaa esteiden poistamiseen, kun taas palkkaeroja samasta työstä voidaan, ainakin periaatteessa, poistaa yksinkertaisesti korjaamalla naisten palkkoja. Nämä seikat huomioon ottaen työssä pyritään selvittämään ainoastaan yhtä työmarkkinasyrjinnän tyyppiä, palkkaeroja samasta työstä. Työtehtävän sisäisetkin palkkaerot voivat olla perusteltuja, jos ne johtuvat työsuorituseroista. Erityisesti katsotaan, että paremmin koulutetut ja kokeneemmat henkilöt tekevät työnsä paremmin tai tehokkaammin, joten heille kuuluu parempi korvaus.

Palkkadiskriminaation arviointi edellyttää, että palkanmuodostusprosessille konstruoidaan tilastollinen malli, josta käy ilmi, miten työn ja henkilön ominaisuuksia kuvaavat tekijät vaikuttavat palkkaan. Lähtökohtana on ajatus, että jos naisten ja miesten välillä ei ole syrjiviä palkkaeroja, miesten ja naisten palkanmuodostukseen vaikuttavat samat tekijät identtisellä tavalla. Jos sitten havaitaan, että joidenkin työn tai henkilön ominaisuuksien vaikutus palkkaan muuttuu sen mukaan, onko kyseessä nais- tai mieshenkilö, tarkoittaa tämä, että kyseisten ominaisuuksien palkitseminen on sukupuolisidonnainen. Miesten ja naisten ominaisuuksien palkitsemiseroilla selittyvä palkkaero antaa tässä tutkimuskehikossa arvion palkkadiskriminaatiosta.

Menetelmällä saadun palkkadiskriminaation arvion tarkkuus ja luotettavuus riippuu siitä, minkälaisen tekijöiden vaikutus palkanmuodostukseen on otettu huomioon ja kuinka tarkasti nämä tekijät mittaavat työtehtävä- ja tuottavuuseroja. Näin ollen tämän tyyppisistä tutkimuksista saatuja arvioita palkkadiskriminaatiosta voidaan aina kritisoida. Mahdolliset heikkoudet eivät kuitenkaan tarkoita, etteikö menetelmä tuota informatiivisia tuloksia.

Selvityksessä käytetty henkilötason aineisto on poikkileikkaus vuodelta 1993 ja paneeli vuosilta 1991 ja 1993. Ne on poimittu Suomen Teollisuuden ja Työnantajain Keskusliiton palkkatilastoista ja edustavat Suomen teollisuudessa työskenteleviä toimihenkilöitä. Aineisto on parhaita saatavilla olevia Suomessa. Se sisältää varsin yksityiskohtaiset tiedot toimialasta, työtehtävän tyyppistä ja vaativuudesta. Tietoa on myös henkilön koulutustasosta ja teknisten ja teollisuustoimihenkilöiden osalta myös työkokemuksesta. Arviot työkokemuksesta tosin osoittautuivat puutteellisiksi.

Toimihenkilöiden keskuudessa naisten keskipalkka on vuonna 1993 keskimäärin 32 % pienempi kuin miesten palkka. Tästä kokonaispalkkaerosta 2/3 - 3/4 johtuu työtehtävä- ja henkilökohtaisten ominaisuuksien eroista ja 1/4-1/3 on arvioitu palkkadiskriminaatiosta johtuvaksi. Luvut riippuvat siitä, onko palkkadiskriminaatio arvioitu keskivertonaisen vai keskivertomiehen kohdalla. Toisin sanoen, keskivertonainen saa noin 10 % pienempää palkkaa kuin samaa työtä tekevä mies. Toisaalta keskivertomies on 13 % paremmin palkattu kuin nainen samankaltaisessa työssä.

Palkkadiskriminaatio toimihenkilöillä vaihtelee mm. toimihenkilötyypeittäin ja toimialoittain. Suhteellisesti eniten palkkadiskriminaatiosta kärsivät ylempät naistoimihenkilöt, eli naisjohtajat ja naisasiantuntijat. Keskimääräinen palkkadiskriminaatio vaihtelee melkoisesti toimialoittain. Arvioitu palkkadiskriminaatio paperiteollisuudessa ja graafisella alalla on lähes 20 %. Toisaalta konsulttien ryhmässä ja mekaanisessa metsäteollisuudessa se on vain 5 %. Palkkadiskriminaatio näyttää olevan tyypillisissä miestehtävissä jossain määrin suurempi kuin tyypillisissä naistehtävissä, mutta ei löytynyt näyttöä siitä, että toimialoittaiset erot diskriminaatioissa olisivat yhteydessä alan miestoimihenkilöiden osuuden kanssa.

Käytetty aineisto ei sisällä tietoja esim. henkilöiden lahjakkuudesta ja ahkeruudesta. Työssä on pohdittu, missä määrin tämänkaltaisten tekijöiden poisjättäminen palkanmuodostuksen tilastollisesta mallintamisesta on vaarallista palkkadiskriminaation arvioinnin kannalta. Vaikka ei ollut vahvoja syitä ajatella, että lahjakkuuden ja ahkeruuden poisjättäminen tilastollisista palkkamalleista heikentäisi ratkaisevasti palkkadiskriminaation analyysia, on tarkasteltu mahdollisuutta ottaa huomioon näiden vaikutusta palkanmuodostukseen paneeliaineiston avulla. Käytetty menetelmä, vakioisten vaikutusten malli, osoittautui varsin ongelmalliseksi, joten tältä osin työ jäi tuloksetta.



# 1 TAVOITTEET, RAJAUKSET JA SISÄLTÖ

## 1.1 TAVOITTEET JA RAJAUKSET

Oikeudenmukaisuus ja tasa-arvoisuus ovat yleisesti hyväksytyjä yhteiskunnallisia tavoitteita. Niiden toteuttaminen koetaan erityisen tärkeäksi päämääräksi juuri työelämässä. Myös kansantaloustieteellinen tutkimus pyrkii vastaamaan yhteiskunnallisen keskustelun esiin nostamiin kysymyksiin ja näin palkkadiskriminaatiotutkimuksesta on muodostunut yksi palkkatutkimuksen keskeisistä osa-alueista. Palkkadiskriminaation kohteena voivat olla useat eri ryhmät, kuten naiset, siirtolaiset tai etniset vähemmistöt. Suomen tämänhetkissä oloissa on naisiin työelämässä kohdistuvaa syrjintää pidetty ajankohtaisimpana, ja se on ollut varsin suuren mielenkiinnon kohteena. Tässä keskitytään yksinomaan sukupuoleen perustuvaan palkkadiskriminaatioon ja palkkadiskriminaatiosta puhuttaessa tarkoitetaan jatkossa pelkästään naisiin kohdistuvaa palkkadiskriminaatiota.

Kansantaloustieteessä palkkadiskriminaatiota tutkitaan teoreettisesti ja empiirisesti. Teoreettisen tutkimuksen tavoitteena on selvittää, minkä mekanismin kautta palkkadiskriminaatio ilmiönä syntyy. Empiirisen tutkimuksen tavoitteena on arvioida (mitata) palkkadiskriminaatiota. Tämä tutkimus on sisällöltään empiirinen. Keskeisin tavoite on palkkadiskriminaation suuruuden arviointi Suomen teollisuuden toimihenkilöiden keskuudessa vuonna 1993. Tarkastellaan myös mahdollisuuksia hyödyntää paneeliaineiston informaatiota palkanmuodostuksen tilastollisessa mallintamisessa ja erityisesti palkkadiskriminaatiotutkimuksessa.

Palkkadiskriminaatiotutkimuksessa täytyy ottaa kantaa siihen, mitä käsite palkkadiskriminaatio tarkoittaa. Kysymys on relevantti, koska sukupuolen vaikutus henkilön työelämään ja palkkaan on hyvin moniulotteinen ilmiö. Keskimäärin miesten ja naisten palkkaero on jopa muutama kymmenen prosenttia. Työmarkkinat ovat vahvasti segregoituja sukupuolen suhteen, eli suuri osa työtehtävistä on selkeästi nais- tai miesvaltaisia, ja naisvaltaiset tehtävät ovat tyypillisesti miesvaltaisia huonommin palkattuja. Tämä ei välttämättä tarkoita,

että tietyn työtehtävän sisällä miesten palkka olisi korkeampi kuin naisten. Saman työtehtävän sisälläkin sukupuolten välinen palkkaero voi olla perusteltu, jos miesten ja naisten osaamisen taso on erilainen. Osaamisero voi johtua mm. koulutuksesta, työkokemuksesta ja työsuhteen kestosta. Mitä palkkadiskriminaatio sitten on? Sukupuoleen perustuvalla palkkadiskriminaatiolla tarkoitetaan tässä tilannetta, jossa naiset ja miehet saavat *samasta työstä* eri suuruista palkkaa. Käsite ”sama työ” on määritelmässä keskeinen. Kaksi henkilöä A ja B tekevät samaa työtä, jos seuraavat ehdot toteutuvat: molemmilla on sama työtehtävä (toimenkuva) ja sama pätevyys (koulutus, työkokemus, ikä). Tämä lista ei ole suinkaan tyhjentävä. Voidaan perustellusti vaatia myös, että A ja B ovat yhtä lahjakkaita, ahkeria ja työkykyisiä. Viimeksi mainituista ominaisuuksista on harvoin tietoa, joten yleensä niiden vaikutusta palkkaan ei erikseen voida analysoida. Kuitenkin katsotaan, että ainakin suurin osa tästä vaikutuksesta tulee välillisesti otetuksi huomioon juuri työtehtävän ja pätevyyden kautta. Myöhemmin pohditaan, mitkä ovat mahdollisia haittoja, jotka koituvat henkilöiden lahjakkuuden ja ahkeruuden huomiotta jättämisestä palkkadiskriminaatiotutkimuksessa.

Tutkimuksessa keskitytään arvioimaan palkkaeroja samasta työstä ja saman työn käsite määritellään niin tiukasti kuin mahdollista analyysissa käytetyn tutkimusaineiston perusteella. Vastaavasti myös palkkadiskriminaatio ymmärretään suppeasti. Palkkadiskriminaatioksi ei tulkita miesten ja naisten välistä palkkaeroa, joka johtuu siitä, että tyypilliset naisvaltaiset tehtävät ovat huonommin palkattuja kuin miesvaltaiset. Analyysia rajoitetaan näin, koska ammatinvalinta ja työtehtävään sijoittuminen ovat monimutkaisen prosessin tulos. Naisilla voi olla perheen takia eri preferenssijärjestys ja huonommat insentiivit kasvattaa inhimillistä pääomaansa. Saattaa olla, että naiset ovat miehiä halukkaampia ottamaan esimerkiksi työpaikan, jossa on säännölliset työajat ja vapaat viikonloput, mutta pienempi palkka. Jos naiset odottavat olevansa työelämässä lyhyemmän aikaa kuin miehet, inhimillisen pääoman teorian tulosten perusteella heidän ei myöskään kannata investoida koulutukseen yhtä paljon kuin miesten. Tämä voi olla syy, miksi naiset sijoittuvat keskimäärin vähemmän vaativiin ja vastaavasti huonommin palkattuihin tehtäviin.<sup>1</sup> Kuitenkin ajatus, että työmarkkinoiden segg-

---

<sup>1</sup> Suomessa nuoret naiset ovat nykyään jopa paremmin koulutettuja kuin nuoret miehet. Tilanne on kuitenkin varsin uusi ja jos verrataan kaikkia työelämässä olevia henkilöitä, havaitaan, että miehillä on

regaatio johtuisi pelkästään preferenssi- ja investointiinsentiivi eroista sukupuolten välillä, on varsin naiivi. Työnantajan päätökset ovat myös keskeisiä työntekijän uran kannalta. Työnantajan haluttomuudelle sijoittaa naisia tiettyihin tehtäviin voi olla erilaisia perusteita. Jos naiset ovat perhesyistä (esim. lastenhoidon vuoksi) keskimäärin miehiä useammin pois työstä ja heidän työuransa jopa katkeaa pitemmäksi ajaksi, työnantaja saattaa olla haluton sijoittamaan naisia työtehtäviin, joissa työntekijää on vaikeaa ja kallista korvata ja joissa poissaolot ovat muutenkin yrityksen toiminnan kannalta erityisen hankalia. Toisaalta naiset ovat tietoisia työnantajan preferensseistä ja päätöksentekoprosessista ja he sopeuttavat niihin ammatinvalintansa ja uraodotuksensa.

Joka tapauksessa naisten sijoittumista työtehtäviin on varsin hankalaa analysoida tarkasti, koska prosessi riippuu osittain naisten omista valinnoista ja osittain työnantajan päätöksistä. Näiden välillä on mutkikas vuorovaikutus. Myös naisten yhteiskunnallinen asema vaikuttaa sekä työnantajan että naistyöntekijöiden valintoihin ja päätöksiin. Ei voida kiistää, etteikö naisten sijoittuminen matalapalkkaisiin tehtäviin heijastaisi jossain määrin naisiin kohdistuvaa syrjintää, mutta työmarkkinoiden segregaation ei voida mitenkään katsoa johtuvan yksinomaan syrjinnästä. On vaikea tehtävä päättää jollain objektiivisella kriteerillä, mikä on syrjinnän osuus tässä prosessissa.

Tutkijat ovat eri mieltä siitä, kuinka yksityiskohtaisesti palkkadiskriminaatiotutkimuksessa tulisi kontrolloida työtehtävän (*occupation*) suhteen, ja sen seurauksena palkkadiskriminaatio ymmärretään toisissa tutkimuksissa laajemmin kuin toisissa. Vastavasti arviot palkkadiskriminaatiosta voivat vaihdella melkoisesti. Mitä tiukemmin analyysissa kontrolloidaan työtehtävän suhteen, sitä pienempi arvio palkkadiskriminaatiosta yleensä saadaan (Gunderson, 1989, 51). Täytyy muistaa, että tieteellisesti eettisessä tutkimuksessa palkkadiskriminaation olemassaolo täytyy todistaa. Kun palkkadiskriminaation käsite määritellään suppeasti, mahdollisuus, ettei löydy riittävää näyttöä palkkadiskriminaation

---

edelleen keskimäärin parempi koulutus. Tässä kontekstissa muutenkin ymmärretään koulutus laajasti, eli tarkoitetaan muodollisen koulutuksen lisäksi myös työelämässä tapahtuvaa kouluttamista (on-the-job training).

olemassaolon puolesta, kasvaa. Toisaalta, jos näyttö on riittävä, palkkadiskriminaation olemassaoloa on vaikea kiistää. Tiukasti määritellyn palkkadiskriminaation arvio voi olla suhteellisen pieni osa keskimääräisestä palkkaerosta sukupuolten välillä, mutta vastaavasti epäily siitä, että se johtuisi jostain muusta kuin naisiin kohdistuvasta syrjinnästä, vähenee. Näistä syistä palkkadiskriminaatio-käsitteellä tarkoitetaan tässä eri palkkaa samasta työstä ja lisäksi "sama työ" määritellään mahdollisimman tiukasti. Määritelmän mukaan palkkaerot, jotka johtuvat työtehtävästä, sen vaativuudesta, henkilön koulutuksesta ja työkokemuksesta (iästä) eivät ole palkkadiskriminaatiota.

Kun palkkadiskriminaation käsite on täsmennetty ja kun otetaan huomioon käytettävissä olevat tutkimusmenetelmät ja tutkimusaineisto, voidaan täsmentää myös työn tavoitteet. Halutaan vastata seuraaviin kysymyksiin:

- onko Suomen teollisuuden toimihenkilöiden keskuudessa palkkaeroja samasta työstä (palkkadiskriminaatiota);
- kuinka suuri on palkkadiskriminaatio ja mistä tekijöistä sen suuruus riippuu;
- mikä on palkkadiskriminaation osuus miesten ja naisten keskipalkkaerosta (vrt. Vartia, 1993, 1-2).

Huomataan, että tilastollisessa analyysissä tarkkuus, jolla saman työn käsitettä voidaan operationalisoida, riippuu tutkimusaineiston sisältämästä informaatiosta, eli siitä, millaisia työtä ja työsuoritusta kuvaavia tekijöitä tutkijalla on käytettävissään, miten tarkasti ne on mitattu jne. Analyysia on aina mahdollista täsmentää, jos saadaan lisää informaatiota, ja vastaavasti mikä tahansa empiirinen tulos voidaan kiistää, väittämällä, että se on aikaansaatu puutteellisen informaation perusteella. Keskeinen kysymys on, vaikuttaako lisäinformaatio oleellisesti tuloksiin. Kuten Vartia toteaa, rakentavan kritiikin täytyy olla ymmärrettävästi argumentoitua, on siis perusteltava selkeästi, millä tavalla analyysistä poisjäänyt informaatio voi vaikuttaa tuloksiin (Vartia, 1995). Tässä pohditaan, voiko puuttuva informaatio henkilökohtaisen lahjakkuuden ja ahkeruuden kaltaisista tekijöistä johtaa virheelliseen palkkadiskriminaatioanalyysiin, ja paneelaineiston avulla yritetään saada informaatiota kyseisistä tekijöistä.

## **1.2 ALAN TUTKIMUS SUOMESSA JA TYÖN YHTEYDET SIIHEN**

Tähän mennessä kattavin<sup>1</sup> palkkadiskriminaatiota käsittelevä empiirinen tutkimus Suomessa on Vartia - Kurjenoja (1992). Siinä tutkitaan palkkadiskriminaatiota metalli- ja metsäteollisuuden suuryrityksissä. Käytetty yksilötason aineisto on poikkileikkausotos Teollisuuden ja Työnantajain Keskusliiton (jatkossa TT) palkkatilastoista. Työntekijät ja toimihenkilöt tutkitaan erikseen. Keskeinen tulos on, että palkkadiskriminaatio on työntekijöiden keskuudessa 4.5 %; toimihenkilöiden osalta se on tyypillisessä naistehtävässä 8.5 % ja tyypillisessä mies-tehtävässä 13 %. Suurin osa miesten ja naisten keskipalkkaerosta johtuu Vartian ym. tutkimuksen perusteella miesten ja naisten työtehtäväeroista ja eroista henkilön työominaisuuksissa (koulutus, ikä yms.).

Palkkadiskriminaatiota Suomessa on selvittänyt myös Asplund ym. (1993) tutkimuksessa, jossa verrataan palkkadiskriminaatiota eri Pohjoismaissa. Suomen osalta tutkimuksessa käytetty aineisto on Työvoimatutkimus (Labour Force Survey) vuodelta 1987, joka kattaa työikäisen väestön (15 - 64 v.). Tutkimuksen perusteella palkkadiskriminaatio on keskimäärin 19 %, mikä on suurin Pohjoismaissa. Tulos on varsin erilainen kuin Vartian, mikä johtunee siitä, ettei Asplundin ym. tutkimuksessa ole yksityiskohtaisesti kontrolloitu työtehtävän suhteen.

Vartian ym. (1992) tutkimuksessa ei ollut yritystä kuvaavia tietoja. Pasanen (1995) tutkii pro gradu-tutkielmassaan, miten yritysindikaattorien lisääminen palkkamalleihin muuttaa Vartian ym. tutkimuksen tuloksia. Yritysindikaattorit olivat kyllä tilastollisesti merkitseviä, mutta eivät vaikuttaneet palkkadiskriminaatioarvioihin ollenkaan.

Myös tätä työtä voidaan pitää Vartian ym. tutkimuksen eräänlaisena jatkona. Työn tavoitteiden asettelussa on otettu huomioon, mitä on jo selvitetty Vartian ym. ja Pasanen

---

<sup>1</sup>Tilanne v. 1995 alkupuolella. Tätä aikaisempia sukupuolten välisiä palkkaeroja Suomessa käsitteleviä tutkimuksia ovat Allen ym. (1990), Brunila (1990) ja Hemmilä (1989). Yhteenveto näiden tutkimusten tuloksista ks. Vartia ym. (1992), 9-12.

tutkimuksissa, ja analyysia on pyritty eräiltä osin laajentamaan ja syventämään. Vartian ym. tutkimus on vaikuttanut taustalla olevaan metodologiaan ja tutkimusmenetelmän valintaan. Myös tietyt tulkinnat ja tekniset ratkaisut ovat peräisin kyseisestä tutkimuksesta. Tämä on luonnollista, sillä tässä käytetty aineisto on rakenteeltaan sama kuin Vartian ym. tutkimuksessa.

### **1.3 TYÖN SISÄLTÖ**

Luvussa 2 esitellään empiirisen palkkadiskriminaatiotutkimuksen menetelmä. Kovarianssi-analyysin avulla kuvataan tilastollisesti palkanmuodostusprosessia, jossa monet tekijät, kuten koulutus, ikä (työkokemus), työtehtävä, toimiala ja sukupuoli, vaikuttavat samanlaisesti palkkojen määräytymiseen. Nollahypoteesina on, ettei palkkadiskriminaatiota ole, eli palkanmuodostus on sukupuolesta riippumaton. Jos nollahypoteesi ei ole voimassa, miesten ja naisten palkanmuodostukseen voivat vaikuttaa eri tekijät tai saman tekijän vaikutus vaihtelee sukupuolen mukaan. Kovarianssianalyysillä pyritään saamaan realistinen kuva palkanmuodostusprosessista ja vastaamaan kysymykseen, millä tavalla palkanmuodostus eroaa sukupuolten välillä. Varsinaista palkkadiskriminaation arviointia varten introduoidaan palkkadiskriminaatiofunktion käsite ja esitetään, millä tavalla miesten ja naisten välinen palkkaero dekomponoidaan kahteen osaan. Toinen on työtehtävistä ja työsuoritukseen vaikuttavista tekijöistä johtuva ja toinen on varsinainen palkkaero samasta työstä eli palkkadiskriminaatio.

Luvussa 3 esitetään tutkimusaineisto, sen poimintatapa ja analyysin kannalta tärkeimmät muuttujat. Luvuissa 4 ja 5 käytetty aineisto on vuoden 1993 TT:n palkkatilastoista poimittu otos toimihenkilöitä, jotka edustavat kaikkia TT:n jäsenyrityksissä työskenteleviä toimihenkilöitä, eli lähes kaikkia Suomen teollisuuden toimihenkilöitä. Luvussa 6 käytetty aineisto on paneeli vuosilta 1991 ja 1993, sekin on poimittu TT:n palkkatilastoista. Paneelin toimihenkilöt edustavat toimihenkilöitä, jotka ovat työskennelleet TT:n jäsenyrityksissä ajanjaksolla 1991-1993.

Luvuissa 4 ja 5 estimoidaan laajoja palkkayhtälöitä ja tutkitaan palkkadiskriminaatiota Suomen teollisuustoimihenkilöiden keskuudessa vuonna 1993. Luvussa 4 kuvataan palkka-

mallin rakentamisprosessia, esitetään estimointitulokset ja muodostetaan palkkadiskriminaatiofunktio. Luvussa 5 arvioidaan palkkadiskriminaatiota ja sen osuus keskipalkkaerosta miesten ja naisten välillä. Palkkadiskriminaatioerot tutkitaan myös toimihenkilöryhmittäin ja toimialoittain, ja esitetään estimoidun palkkadiskriminaation jakauma tutkimusaineistossa.

Palkanmuodostukseen vaikuttavat muun muassa lahjakkuuden ja ahkeruuden kaltaiset vaikeasti mitattavat tekijät. Koska niistä ei yleensä ole tietoa, niiden vaikutusta palkkaan ei voida ottaa huomioon poikkileikkausanalyysissa. Luvussa 6.1 pohditaan, millä tavalla ja missä määrin tällaiset tekijät voivat aiheuttaa ongelmia palkkadiskriminaation arvioinnissa, mikäli niiden vaikutusta palkanmuodostukseen ei oteta huomioon palkkamalleissa. Todetaan, ettei ole perusteltua ajatella ei-havaittavien tekijöiden kuten lahjakkuuden ja ahkeruuden olevan palkkadiskriminaatiotutkimuksessa päällimmäinen ongelma. Toisaalta ilman empiiristä näyttöä on vaikea väittää, että ne olisivat täysin merkityksettömiä palkkadiskriminaation arvioinnin kannalta. Luvussa 6.2 esitetään, miten ei-havaittavien tekijöiden vaikutusta palkanmuodostukseen voidaan periaatteessa mallintaa paneelitutkimuksessa. Metodi ei kuitenkaan ole luonteva palkkadiskriminaatiotutkimuksessa. Ainakin tässä tapauksessa identifioitavuus- ja multikollineaarisuusongelmien ja aineiston ei-kokeellisen luonteen vuoksi ajaudutaan tilanteeseen, jossa normaali tilastollinen päättely ei onnistu. Tämä käy ilmi luvuissa 6.3 ja 6.4.

## 2 TUTKIMUSMENETELMÄ

### 2.1 LÄHTÖKOHDAT

Luvussa 1.1 todettiin, että palkkadiskriminaatiolla tarkoitetaan tässä palkkaeroa samasta työstä. Empiirisissä tutkimuksissa saman työn käsite täytyy operationalisoida. Analyysi perustuu inhimillisen pääoman teoriaan. Teorian lähtökohta on neoklassisen työn talousteorian, jonka mukaan täydellisillä työmarkkinoilla palkkaerot johtuvat yksinomaan työntekijöiden rajatuottavuuseroista. Inhimillisen pääoman teoria katsoo, että koulutus (sekä muodollinen koulutus (*schooling*) että työpaikalla tapahtuva kouluttaminen (*on-the-job training*)) on prosessi, jonka kautta henkilöt kasvattavat osaamistaan, eli inhimillistä pääomaansa. Inhimillisen pääoman varannon kasvu parantaa työn rajatuottavuutta ja täydellisillä työmarkkinoilla tämä implikoi myös korkeampaa palkkaa. Empiirisissä tutkimuksissa inhimillisen pääoman tasoa arvioidaan sellaisilla muuttujilla kuten koulutus, työkokemus (ikä) ja työsuhteen kesto. Vastaavasti palkanmuodostusta mallinnetaan muodostamalla regressiomalli, jossa logaritminen palkka selitetään inhimillistä pääomaa kuvaavilla muuttujilla.<sup>1</sup> Inhimillistä pääomaa kuvaavat muuttujat eivät yksin kuvaa riittävän hyvin palkanmuodostusprosessia ja palkkaeroja, joten empiirisissä palkkamalleissa on yleensä muitakin selittäjiä, jotka kuvaavat työtehtävää, toimialaa, henkilön perhetaustaa yms. Työtehtävää ja toimialaa kuvaavien muuttujien mukanaolo palkanmuodostuksen kuvaamisessa on tärkeää tässä tutkimuksessa, koska tavoitteena on arvioida miesten ja naisten palkkaerot samasta työstä.

Palkkadiskriminaatiotutkimuksessa keskeisin palkanmuodostuksen selittäjä on sukupuoli. Jos palkkadiskriminaatiota ei ole, palkanmuodostus on sukupuolesta riippumaton, kun inhimillistä pääomaa kuvaavien muuttujien ja työtehtävän ominaispiirteiden vaikutus palkanmuodostukseen on otettu huomioon. Tällöin odotettu palkka tietyn henkilön kohdalla riip-

---

<sup>1</sup> Mincer (1974) perustelee yksinkertaisen lineaarisen relaation henkilön logaritminen ansioiden (palkan) ja inhimillisen pääoman tasoa kuvaavien muuttujien (koulutusvuodet, kokemusvuodet, kokemusvuodet potenssiin kaksi) välillä. Ks. Asplund, 1993a, 11-17.



puu ainoastaan työn luonteesta ja henkilön inhimillisestä pääomasta. Jos sitä vastoin todetaan eroja miesten ja naisten palkanmuodostuksessa, odotettu palkka on ainakin tietyissä tilanteissa sukupuolesta riippuvainen, eli työmarkkinoilla esiintyy palkkadiskriminaatiota.

Palkanmuodostus mallinnetaan tässä kovarianssianalyysillä, joka on yleisesti käytetty menetelmä inhimillisen pääoman teoriaan perustuvassa palkkatutkimuksessa. Kovarianssianalyysillä pyritään saamaan realistinen kuva palkanmuodostuksesta ja tutkitaan, millä tavalla sukupuoli vaikuttaa siihen. Kovarianssianalyysimenetelmä esitetään yksityiskohtaisemmin luvussa 2.1. Kun palkanmuodostuksesta on saatu luotettava kuva, voidaan palkkamallien perusteella laskea arviot palkkadiskriminaatiolle. Palkkadiskriminaation arviointia käsitellään kohdassa 2.3.

## **2.2 KOVARIANSSIANALYYSIN ESITTELY PALKKADISKRIMINAATIOTUTKIMUKSEN YHTEYDESSÄ**

Kovarianssianalyysi on kahden muun tilastollisen menetelmän, varianssi- ja regressioanalyysin yhdistelmä. Sitä voidaan luonnehtia indikaattorimuuttujilla täydennetyksi regressioanalyysiksi. Indikaattorimuuttuja on muuttuja, joka saa arvon yksi, kun havaintoaineiston i:nnen havainto toteuttaa tietyn ehdon ja arvon nolla muissa tapauksissa. Esimerkiksi voidaan konstruoida indikaattorimuuttuja  $N$  niin, että  $N=1$ , kun henkilö on nainen ja  $N=0$ , kun kyseessä on mies, jolloin  $N$  on *naisindikaattori*. Kovarianssianalyysin soveltaminen palkkatutkimuksessa tarkoittaa, että palkkojen vaihtelu henkilöstä toiseen kuvataan kvalitatiivisilla tekijöillä (esim. sukupuoli, yritys, toimiala, ammattiliitto) ja kvantitatiivisilla muuttujilla (ikä, koulutusvuodet) sekä niiden avulla muodostetuilla tulomuotoisilla muuttujilla, interaktioilla. Kovarianssianalyysin tuloksena saadaan malleja, joissa kiinnostava informaatio on selkeässä, kompaktissa muodossa.

Kovarianssianalyysin käyttöä palkkamallien rakentamisessa havainnollistetaan aluksi esimerkin avulla. Oletetaan alustavasti, että toimihenkilöiden (logaritmiset) kuukausipalkat ( $P$ ) voidaan riittävän tarkasti selittää sukupuolella ( $N =$  naisindikaattori), koulutuksella ( $K$ ),

iällä (I), työkokemuksella (TK) ja toimihenkilöryhmällä. TT:n palkkatilastoissa on kolme toimihenkilöryhmää: ylemmät, tekniset ja konttoritoimihenkilöt. Tarvitaan kaksi toimihenkilöryhmäindikaattoria - esimerkiksi ylempien toimihenkilöiden indikaattori (YTH) ja teknisten toimihenkilöiden indikaattori (TTH), jolloin konttoritoimihenkilöt muodostavat ns. *vertailuryhmän*. Malli oletetaan lineaariseksi, eli ajatellaan, että jokaisen toimihenkilön kohdalla palkanmuodostusta kuvaa *palkkayhtälö*

$$(2.1) \quad P = \beta_0 + \beta_1 K + \beta_2 I + \beta_3 TK + \beta_4 YTH + \beta_5 TTH + \beta_6 N + \varepsilon,$$

jossa  $\varepsilon$  on jäännöstermi, eli mallin ulkopuolelle jätetyt tekijät palkanmuodostuksessa (tässä esimerkiksi työtehtävä, toimiala tai vaikkapa asenteet). Jäännöstermistä yleensä oletetaan, että se on jokaisen toimihenkilön kohdalla normaalijakautunut satunnaismuuttuja, jonka saama arvo on keskimäärin nolla ja eri toimihenkilöihin liittyvät jäännöstermit ovat keskenään korreloimattomia. Yhtälöä (2.1) voidaan kutsua toimihenkilön *työnarvostusfunktioksi* (Vartia 1992, 24). Tehtävänä on estimoida tuntematon *kiinteä* (ei-satunnainen) parametrivektori  $\beta = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6)$ , jonka elementit kertovat, kuinka suuri on eri mallin selittäjien (iän, koulutuksen, toimihenkilöryhmän jne.) vaikutus palkanmuodostuksessa. Olkoon  $\mathbf{b} = (b_0, b_1, b_2, b_3, b_4, b_5, b_6)$  sopivan estimointimenetelmän (yleensä pienimmän neliösumman PNS) antama  $\beta$ :n estimaatti. Tilastollisten testien avulla voidaan selvittää, onko joku selittäjä tärkeä palkanmuodostuksen kannalta. Oletetaan, että malli (2.1) on hyvin spesifioitu, eli se kuvaa riittävän tarkasti todellista palkanmuodostusta toimihenkilöillä. Tällöin, jos esimerkiksi  $b_5$  ei poikkeaa merkitsevästi nolasta, voidaan päätellä, että ylempien toimihenkilöiden indikaattori ei selitä palkanmuodostusta, *kun muiden mallin selittäjien vaikutus on otettu huomioon*. Tämä ei suinkaan tarkoita, etteikö ylempien toimihenkilöiden palkka olisi keskimäärin erilainen kuin vertailuryhmän (konttoritoimihenkilöiden) palkka, vaan että mallin muut selittäjät jo selittävät tämän palkkaeron kokonaisuudessaan.

Malli (2.1) ei luultavasti ole hyvin spesifioitu, ei pelkästään siitä puuttuvien tärkeiden selittäjien (esim. työtehtävä ja toimiala) vuoksi, vaan erityisesti siksi, että sen rakenne on kovin rajoittava palkkadiskriminaation tutkimisen kannalta. Malli sallii vain sellaisia eroja naisten

ja miesten palkanmuodostuksessa, jotka voidaan kuvata poikkeavilla vakiotermeillä. Jos tarkastelut suoritetaan spesifikaation (2.1) puitteissa, tarkoittaa se, että muiden selittäjien vaikutus oletetaan alusta asti samanlaiseksi sukupuolesta riippumatta, mikä on mielivaltaista ja osoittautuu virheelliseksi lähtökohdaksi. Palkkadiskriminaatioanalyysissä tehtävänä on juuri tutkia ja kuvata, miten miesten ja naisten palkanmuodostukset eroavat. Mallia (2.1) hieman joustavampi olisi esimerkiksi sellainen malli, jossa selittäjäksi on otettu mukaan koulutuksen ja naisindikaattorin interaktiotermin  $N \cdot K$ . Sen avulla voidaan tutkia, onko koulutuksen vaikutus samanlainen miesten ja naisten kohdalla. Tällä interaktiolla täydennetty palkkayhtälö on

$$(2.2) \quad P = \beta_0 + \beta_1 K + \beta_2 I + \beta_3 TK + \beta_4 YTH + \beta_5 TTH + \beta_6 N + \beta_7 N \cdot K + \varepsilon .$$

Parametri  $\beta_7$  mallissa (2.2) kertoo, paljonko koulutuksen vaikutus palkkaan poikkeaa naisilla verrattuna miehiin. Samalla tavalla voidaan edelleen laajentaa mallia (2.3) siten, että otetaan siihen selittäjiksi kaikki naisindikaattorin ja muiden mallin (2.1) selittäjien interaktiot. Näin saadaan varsin mutkikkaalta vaikuttava malli, joka on muotoa

$$(2.3) \quad P = \beta_0 + \beta_1 K + \beta_2 I + \beta_3 TK + \beta_4 YTH + \beta_5 TTH + \beta_6 N + \beta_7 N \cdot K + \beta_8 N \cdot I + \\ + \beta_9 N \cdot TK + \beta_{10} N \cdot YTH + \beta_{11} N \cdot TTH + \varepsilon .$$

Malli (2.3) on itse asiassa yhtäpitävä sen kanssa, että miehille ja naisille konstruoidaan eri mallit. Tämän seikan havainnollistamiseksi yksinkertaistetaan tilannetta. Ajatellaan, että kaikki toimihenkilöiden palkat riippuvat ainoastaan koulutusvuosista ja sukupuolesta ja kirjoitetaan palkkayhtälöt tässä tapauksessa erikseen miehille ja naisille<sup>1</sup>:

---

<sup>1</sup> Yksinkertaisuuden vuoksi oletetaan, että käytetyt selittäjät kuvaavat yhtä hyvin sekä miesten että naisten palkanmuodostusta, jolloin virhetermit miesten ja naisten malleissa ovat samoja (samalla tavalla jakautuneita) ja niiden varianssi on sama. Näin ei tietenkään tarvitse olla. Virhetermien varianssin erilaisuus miesten ja naisten malleissa ei kumoa jatkossa esitettäviä perusteluja, jotka liittyvät mallin rakentamiseen tekniikkaan. Jos virhetermin varianssi on erilainen miesten ja naisten malleissa, yhteisen mallin virhetermi on hete-

$$(2.4) \quad P_m(K) = \beta_0 + \beta_1 K + \varepsilon \quad \text{miesten palkkayhtälö}$$

$$(2.5) \quad P_n(K) = \gamma_0 + \gamma_1 K + \varepsilon \quad \text{naisten palkkayhtälö .}$$

Määritellään parametrit  $\beta_2 = (\gamma_0 - \beta_0)$  ja  $\beta_3 = (\gamma_1 - \beta_1)$ . Tästä seuraa, että  $\gamma_0 = (\beta_0 + \beta_2)$  ja vastaavasti  $\gamma_1 = (\beta_1 + \beta_3)$ . Tällöin naisten palkkayhtälö voidaan kirjoittaa muodossa

$$(2.6) \quad \begin{aligned} P_n(K) &= \gamma_0 + \gamma_1 K + \varepsilon = (\beta_0 + \beta_2) + (\beta_1 + \beta_3)K + \varepsilon = \\ &= P_m(K) + (\beta_2 + \beta_3 K) . \end{aligned}$$

Tämä yhtälö ilmoittaa naistoimihenkilön palkkayhtälön miestoimihenkilön palkkayhtälön avulla. Se kertoo, että naistoimihenkilön palkkayhtälön voidaan ajatella koostuvan kahdesta osasta: koulutustasoltaan samanlaisen miestoimihenkilön palkkayhtälöstä  $P_m(K)$  ja “naislisästä” (tai “naisvähennyksestä”)  $(\beta_2 + \beta_3 K)$ . Tästä seuraa, että i:nnen toimihenkilön palkkayhtälö on

$$(2.7) \quad P(K_i) = \beta_0 + \beta_1 K_i + \beta_2 N_i + \beta_3 (N^*K)_i + \varepsilon_i .$$

Jos i:nnen toimihenkilö on mies,  $N_i = 0$  ja  $(N^*K)_i = 0$ , joten (2.7) on yhtäpitävä esityksen (2.4) kanssa. Jos kyseinen toimihenkilö on nainen,  $N_i = 1$  ja  $(N^*K)_i = K_i$ , joten (2.7) on tässä tapauksessa sama kuin (2.6).

Aivan samalla tavalla kuten edellä esitettyssä esimerkissä, voidaan menetellä, kun palkkayhtälössä on monta selittäjää. Erityisesti (2.3) on yhtäpitävä seuraavan kahden palkkayhtälön kanssa

$$(2.8) \quad P = \beta_0 + \beta_1 K + \beta_2 I + \beta_3 TK + \beta_4 YTH + \beta_5 TTH + \varepsilon, \quad \text{jos henkilö on mies}$$

---

roskedastinen. Tällä seikalla on merkitystä ainoastaan mallin estimoinnin ja ennen kaikkea tilastollisen testauksen kannalta.

$$(2.9) \quad P = \gamma_0 + \gamma_1 K + \gamma_2 I + \gamma_3 TK + \gamma_4 YTH + \gamma_5 TTH + \varepsilon, \text{ jos henkilö on nainen.}$$

Käytännössä tutkimusaineistoa luokitellaan (ryhmitellään) paitsi sukupuolen mukaan myös muilla perusteilla. Jatkossa nähdään, että esimerkiksi toimihenkilöryhmän mukainen luokitus on tärkeä. Sopivilla interaktioilla voidaan ilmaista esimerkiksi eroja iän vaikutuksessa palkkaan toimihenkilöryhmittäin. Tarvittaessa voidaan myös muodostaa ns. toisen asteen interaktioita. Esimerkiksi, (2.3) merkintöjä käyttäen, toisen asteen interaktion ( $I*N*YTH$ ) avulla mallinnetaan poikkeavaa iän vaikutusta naispuolisten ylempien toimihenkilöiden ryhmässä. Korkeamman (esim. neljännen) asteen interaktioiden systemaattinen tutkiminen ei ole kiinnostavaa eikä oikein mahdollistakaan niiden lukuisuuden vuoksi. Esimerkit osoittavat, että kovarianssianalyysi tarjoaa hyvin yleisen ja monipuolisen menetelmän palkkayhtälöiden rakentamiselle.

Yllä esitettiin, miten ns. vapaat mallit miehille ja naisille voidaan interaktioteknikalla yhdistää yhteiseen malliin. Yleensä palkkadiskriminaatiotutkimuksessa spesifoidaan ja estimoidaan naisille ja miehille erilliset palkkayhtälöt. Erillisten palkkayhtälöiden estimointi johtaa siihen, ettei tule aivan selväksi, mikä on yhteistä ja mikä erilaista miesten ja naisten palkanmuodostuksessa. Vartia korostaa, että palkkadiskriminaatiotutkimuksessa lähtökohta, ns. nollahypoteesi, on, että naisten ja miesten palkanmuodostus on samanlainen. Tällöin palkkamallissa esiintyvien sukupuolten palkanmuodostuksen erojen tulee olla tilastollisesti todistettuja (Vartia, 1993, 3). Tämän metodologisen lähtökohdan mukaan miehille ja naisille käytetään erillisiä palkkamalleja korkeintaan alustavissa tarkasteluissa, jotta saataisiin käsitys siitä, mitkä ovat potentiaalisesti merkitsevät erot miesten ja naisten palkanmuodostuksessa. Jatkoanalyysissä siirrytään yhteiseen palkkamalliin, jossa selittäjinä ovat muun muassa joko kaikki<sup>1</sup> tai ainakin kiinnostavimmat naisindikaattorin ja muiden selittäjien interaktiot. Sen jälkeen etsitään intensiivisesti tilastollista näyttöä siitä, minkä tekijöiden suhteen naisten ja miesten palkanmuodostukset eroavat. Tarkoituksena on saada aikaan lopullinen malli, jossa ovat mukana selittäjinä vain naisindikaattorin interaktiot niiden selittäjien kanssa,

---

<sup>1</sup>Jos rakennetaan yhteinen palkkamalli, jossa ovat kaikki naisindikaattorin ja muiden selittäjien interaktiot, ei tietenkään tarvitse välttämättä lainkaan rakentaa ja estimoida miehille ja naisille erillisiä palkkamalleja.

joiden erilaisesta vaikutuksesta naisten ja miesten palkanmuodostukseen on riittävä tilastollinen näyttö. Tässä yllä kuvattua menetelmää on käytetty palkanmuodostuksen ja sukupuolten välisten palkkaerojen mallintamisessa. Se on teknisesti vaativampi kuin erillisten palkkamallien rakentaminen ja sen yhteydessä joudutaan suorittamaan monivaiheinen mallin valinta. Mallin valinta tilastollisessa analyysissä on aina jossain määrin subjektiivista, koska on tilanteita, joissa ei ole yksiselitteistä, pitäisikö ottaa vai olla ottamatta malliin tietty selittäjä.

Kovarianssianalyysin tekniikkaa pyrittiin edellä havainnollistamaan suhteellisen yksinkertaisten esimerkkien avulla. Erityisesti haluttiin korostaa, miten kovarianssianalyysillä voidaan kuvata palkanmuodostusta ja tutkia siinä esiintyviä eroavuuksia sukupuolten välillä. Vaikka todelliset palkkamallit sisältävät kymmeniä selittäjiä, ei se tuo enää uusia metodologisia ongelmia.

### **2.3 PALKKAMALLIEN LOGARITMISESTA MUODOSTA**

Empiirisessä palkkatutkimuksessa selitettävä muuttuja on lähes aina palkan luonnollinen logaritmi, tai joku vastaava logaritmoitu suure. Tässä selitettävä muuttuja on henkilön palkan ja yleiskeskipalkan logaritmoitu suhde. Usein myös osa selittäjistä on logaritmisessa muodossa. Logaritmoinnilla on erittäin vahvat talousteoreettiset, tilastolliset ja tulkinnalliset perustelut.<sup>1</sup> Jäljempänä esitettävien empiiristen tulosten oikean tulkitsemisen kannalta seuraavat huomautukset ovat keskeisiä (ks. tarkemmin Törnqvist ym. 1985 ja Vartia ym. 1992, 26-27).

Kun selitettävä muuttuja on logaritminen palkka, selittäjien kertoimet ilmoittavat, mikä on palkan suhteellinen muutos, kun tietyn selittäjän arvo muuttuu yhdellä yksiköllä muiden selittäjien ollessa muuttumattomia (ns. *ceteris paribus* - oletuksen vallitessa). Esimerkiksi

---

<sup>1</sup> Mincer (1974) perustelee palkkayhtälön puolilogaritmista muotoa lähtien inhimillisen pääoman teoriasta. Tilastollisen analyysin kannalta palkat logaritmoidaan, koska logaritmoidun palkan jakauma on lähempänä normaalijakaumaa, kun taas ei-logaritmoidun (metrisen) palkan jakauma on vino.

jos palkkayhtälössä selitettävä muuttuja on logaritminen palkka ja yksi mallin selittäjistä on ikä, jonka kerroin on 0.005, tarkoittaa tämä, että palkka nousee *ceteris paribus* 0.5 %, kun ikävuodet kasvavat yhdellä yksiköllä (tässä tapauksessa yksikkö on vuosi). Jos myös selittäjä on logaritmisessa muodossa, sen kerroin on *jousto*: Se ilmoittaa (likimain) montako prosenttia palkka muuttuu, kun selittäjän arvo muuttuu, *ceteris paribus*, yhdellä prosentilla.

Logaritminen palkkaero,  $\ln(w_1) - \ln(w_0)$ , ilmoittaa kahden henkilön välisen suhteellisen palkkaeron tai saman henkilön palkassa tapahtuneen suhteellisen muutoksen, log-muutoksen. Sadalla kerrottuna eli prosenteiksi muunnettuna log-muutos kertoo suhteellisen muutoksen logaritmisissa prosenteissa (log-%). Log-prosenttimuutoksen merkitys on suurin piirtein sama kuin tavallisen prosenttimuutoksen. Kun kyseessä ovat kohtuullisen pienet muutokset, logaritmiset ja tavalliset prosenttimuutokset ovat käytännöllisesti katsoen samoja. Log-muutoksista siirrytään tavallisiin suhteellisiin muutoksiin seuraavalla tavalla. Merkitään  $\ell$ llä log-muutosta  $w_0$ :sta  $w_1$ :een

$$(2.10) \quad \ell = \ln(w_1) - \ln(w_0) = \ln(w_1 / w_0) .$$

Tavallinen suhteellinen muutos  $w_0$ :sta  $w_1$ :een, joka merkitään  $p$ :llä on

$$(2.11) \quad p = (w_1 - w_0) / w_0 = \exp(\ell) - 1 .$$

Log-prosenttimuutos on  $100 \cdot \ell$  ja tavallinen prosenttimuutos on  $100 \cdot p$ . Jatkossa tulokset esitetään pääsääntöisesti logaritmisissa prosenteissa. Logaritmisten prosenttien käytön etuna on, että toisin kuin tavallisia prosenttilukuja käytettäessä, on samantekevää, mihin absoluuttista muutosta tai eroa verrataan. Esimerkiksi 10 on 100 % suurempi kuin 5, mutta 5 on vain 50 % pienempi kuin 10. Log-prosenteissa ero on molempiin suuntiin  $100 \cdot \ln(2) = 69$  log-%. Toisaalta logaritmiset prosentit ovat additiivisia, eli esim. kaksi peräkkäistä 5 log-% muutosta on yhteensä 10 log-% muutos. Tämä seuraa suoraan logaritmin keskeisestä ominaisuudesta  $\ln(ab) = \ln(a) + \ln(b)$ . Tavallisia prosenttimuutoksia ei voi laskea yhteen tällä tavalla.

## 2.4 PALKKADISKRIMINAATION ARVIOINTI

Edellä esitettiin palkkayhtälön (työnarvostusfunktion) estimoinnin yleiset periaatteet kovarianssianalyysin kehikossa. Varsinaisena kiinnostuksen kohteena on kuitenkin palkkadiskriminaation mittaaminen tutkittavassa populaatiossa. Tähän tarkoitukseen Vartia määrittelee ns. *palkkadiskriminaatiofunktion* (vrt. Vartia 1993, 2). Olkoon  $F_m(\mathbf{X}, \text{Mies})$  miesten palkkayhtälö ja  $F_n(\mathbf{X}, \text{Nainen})$  naisten palkkayhtälö, jossa  $\mathbf{X}$ :llä merkitään palkkaan vaikuttavia tekijöitä sukupuolta lukuunottamatta. Palkkadiskriminaatiofunktio (PDF) on siten

$$(2.12) \quad \text{PDF}(\mathbf{X}) = F_m(\mathbf{X}) - F_n(\mathbf{X}) .$$

Kaavassa (2.12)  $F_m(\mathbf{X})$  on henkilön odotettu palkka, jos hän on mies. Hänen kohdallaan palkkaan vaikuttavat tekijät sukupuolta lukuunottamatta ovat  $\mathbf{X}$ . Vastaavasti  $F_n(\mathbf{X})$  on kyseisen henkilön odotettu palkka, jos hän on nainen. Näin ollen  $\text{PDF}(\mathbf{X})$  on odotettu sukupuolesta johtuva palkkaero vakioidussa työtilanteessa  $\mathbf{X}$ , eli odotettu palkkadiskriminaatio työtilanteessa  $\mathbf{X}$ . Määritelmässä on epäoleellista, kumpaa sukupuolta henkilö sattuu olemaan. Jos työtilanteessa  $\mathbf{X}$  oleva henkilö on todellisuudessa mies, palkkadiskriminaatiofunktio kertoo, paljonko hänen palkkansa vähenisi, jos hän olisi nainen. Vastaavasti jos henkilö on nainen,  $\text{PDF}(\mathbf{X})$  kertoo, minkä verran hänen palkkansa kasvaisi, jos hän olisi mies. Palkkadiskriminaatiofunktion avulla voidaan arvioida odotettu palkkadiskriminaatio sekä aggregaattitasolla, esimerkiksi naisille tai miehille tyypillisessä tehtävässä tai keskimäärin toimialoittain, ja haluttaessa myös henkilötasolla, eli siinä konkreettisessa työtilanteessa, jossa annettu henkilö on. Itse asiassa voidaan osoittaa, että esim. palkkadiskriminaatio tyypillisessä naistehtävässä (tyypillisen naisen kohdalla) on ainakin likimain keskiarvo palkkadiskriminaation henkilökohtaisista arvioista naisten osajoukossa (ks. Vartia ym. 1992, 82-83 alaviite). Kun palkkayhtälöissä selitettävä muuttuja on logaritminen palkka, on luontevaa siirtyä myös palkkadiskriminaatiofunktiossa logaritmiseen versioon. Logaritminen palkkadiskriminaatiofunktio,  $\text{LPDF}(\mathbf{X})$  ilmoittaa odotetun suhteellisen palkkadiskriminaation, eli se vastaa kysymykseen, kuinka suuri on keskimäärin miesten ja naisten suhteellinen



palkkaero tietyssä työtehtävässä ja henkilön työominaispiirteiden kohdalla. Palkkadiskriminaatiofunktion arvo tilanteessa  $X$  sadalla kerrottuna on siis (odotettu) palkkadiskriminaatio kyseisessä tilanteessa log-prosenteissa ilmaistuna.

Jos tietyn tekijän vaikutus palkkaan on sama sukupuolesta riippumatta, sen kerroin palkkadiskriminaatiofunktiossa on nolla. Estimoidussa palkkadiskriminaatiofunktiossa nolasta poikkeavia kertoimia saavat vain tekijät, joiden vaikutus on tutkimusaineiston perusteella todistettavasti erilainen sukupuolten välillä. Estimoitu palkkadiskriminaatiofunktio muodostetaan suoraan estimoidusta palkkayhtälöstä. Esimerkin vuoksi oletetaan, että on rakennettu ja estimoitu palkkamalli ja ainoat todistetut erot miesten ja naisten palkanmuodostuksessa ovat seuraavat: Naisten palkkojen lähtötaso poikkeaa miesten palkkojen lähtötasosta (palkkayhtälöiden vakiotermit ovat erilaisia), ja myös iän vaikutus palkkaan on erilainen miesten ja naisten kohdalla. Tällöin estimoitu palkkadiskriminaatiofunktio on

$$(2.13) \quad \text{PDF}(X) = \text{PDF}(IK\ddot{A}) = -a - bIK\ddot{A} .$$

Kaavassa (2.13)  $a$  on naisindikaattorin estimoitu kerroin palkkayhtälössä ja  $b$  on naisindikaattorin ja iän ( $N*IK\ddot{A}$ ) interaktion estimoitu kerroin. Palkkadiskriminaatiofunktio on selvästi vakio vain siinä erikoisessa tapauksessa, jossa miesten ja naisten palkkayhtälöt eroavat ainoastaan vakiotermin suhteen. Vartian ym. (1992) mukaan logaritminen palkkadiskriminaatiofunktio on vakio, 0.046 (4.6 log-%) metalli- ja metsäteollisuuden suuryritysten työntekijöillä v. 1992.

Yleisesti ottaen voidaan odottaa, että estimoitu palkkadiskriminaatiofunktio on paljon vähemmän argumentteja sisältävä funktio kuin estimoitu palkkayhtälö, sillä miesten ja naisten palkanmuodostuksen odotetaan olevan todistettavasti erilainen vain joidenkin palkanmuodostukseen vaikuttavien tekijöiden suhteen. Tekijät, joiden suhteen miesten ja naisten palkanmuodostukset ovat samanlaisia, supistuvat pois palkkadiskriminaatiofunktioista.

Palkkadiskriminaatiotutkimuksen keskeisimpiä kysymyksiä on, kuinka suuri osa miesten ja naisten keskimääräisestä palkkaerosta johtuu miesten ja naisten työtehtäväeroista, eroista koulutuksessa, iässä (työkokemuksessa) tms. ja mikä on varsinaisen palkkadiskriminaation osuus. Konstruoidaan hypoteettinen henkilö, jota kutsutaan tyypilliseksi (keskiverto) mieheksi. Tyypillinen mies on henkilö, jonka kohdalla palkkaan vaikuttavat tekijät saavat yleensä<sup>1</sup> arvoksi kyseisten tekijöiden keskiarvot miesten osajoukossa. Vastaavalla tavalla määritellään tyypillinen nainen. Olkoon  $\bar{X}_m$  selittäjävektorin arvo otoksen tyypillisen miehen kohdalla ja  $\bar{X}_n$  selittäjävektorin arvo otoksen tyypillisen naisen kohdalla. Odotettu kokonaispalkkaero tyypillisen naisen ja tyypillisen miehen välillä voidaan esittää joko muodossa

$$(2.14) \quad F_m(\bar{X}_m) - F_n(\bar{X}_n) = F_m(\bar{X}_n) - F_n(\bar{X}_n) + F_m(\bar{X}_m) - F_m(\bar{X}_n) = \\ \text{PDF}(\bar{X}_n) + F_m(\bar{X}_m) - F_m(\bar{X}_n)$$

tai muodossa

$$(2.15) \quad F_m(\bar{X}_m) - F_n(\bar{X}_n) = F_m(\bar{X}_m) - F_n(\bar{X}_m) - F_n(\bar{X}_n) + F_n(\bar{X}_m) = \\ \text{PDF}(\bar{X}_m) + F_n(\bar{X}_m) - F_n(\bar{X}_n).$$

Kaavan (2.14) mukaan odotettu palkkaero keskivertonaisen ja keskivertomiehen välillä koostuu kahdesta osasta:

- Odotetusta palkkadiskriminaatiosta **tyypillisen naisen kohdalla**,  $\text{PDF}(\bar{X}_n)$ ;
- Tyypillisen naisen ja tyypillisen miehen työtehtävästä ja muista palkkaan vaikuttavista ominaisuuksista johtuvasta odotetusta erosta **miesten työnarvostusfunktiolla mitattuna**,  $F_m(\bar{X}_m) - F_m(\bar{X}_n)$ .

---

<sup>1</sup>Tietyissä tilanteissa on luontevampaa, jos muuttujan arvoksi tyypillisen naisen (tai miehen) kohdalla valitaan muu kuin kyseisen muuttujan otoskeskiarvo. Esimerkiksi tämän tutkimuksen malleissa on selittäjänä sekä (lineaarinen) ikä että iän neliöjuuri. Jos naisten keski-ikä on otoksessa 36 vuotta, tuntuu luonnolliselta, että tyypillisen naisen kohdalla iän neliöjuuri saa arvon 6. Kuitenkaan tämän muuttujan keskiarvo otoksessa ei ole 6 (ero on yleensä pieni).

Vastaavasti kaavan (2.15) mukaan odotettu palkkaero tyypillisen naisen ja tyypillisen miehen välillä koostuu:

- Odotetusta palkkadiskriminaatiosta **tyypillisen miehen kohdalla**;
- Tyypillisen naisen ja tyypillisen miehen työtehtävästä ja muista palkkaan vaikuttavista ominaisuuksista johtuvasta odotetusta erosta **naisten työnarvostusfunktiolla mitattuna**.

(2.14) ja (2.15) mukaisen palkkaeron dekomponointi on alun perin Oaxacan (vrt. Oaxaca 1973, 696-697) esittämä. Kumpi yhtälöistä, (2.14) vai (2.15), on "oikea" tapa dekomponoida odotettu keskimääräinen palkkaero, on indekseoteoreettinen ongelma.<sup>1</sup> Tutkimuksessa kannattaa raportoida molemmat.

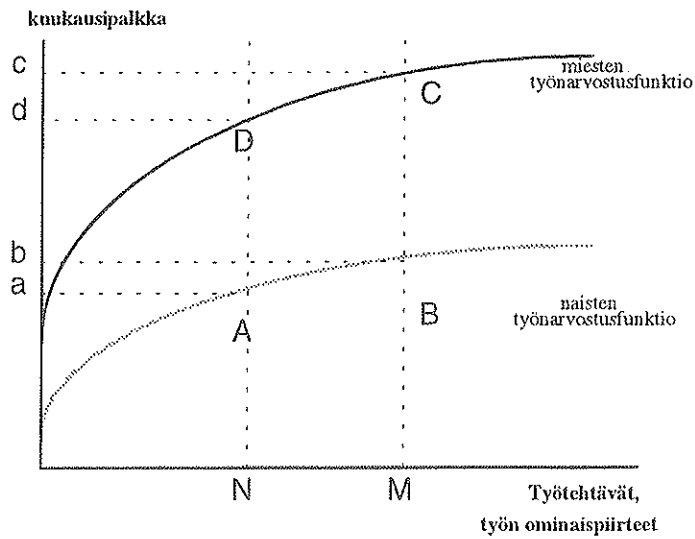
Olkoon  $\bar{P}_m$  miesten ja  $\bar{P}_n$  naisten keskipalkka otoksessa. Jos palkkamalli on estimoitu PNS-menetelmällä, miesten ja naisten kohdalla on erillinen vakiotermin (mallissa on vakiotermin ja pelkkä naisindikaattori on myös mallin selittäjä) ja tämän lisäksi malli on lineaarinen kaikkien selittäjien suhteen (jolloin siis tyypillisen naisen ja tyypillisen miehen kohdalla kunkin selittäjän arvo on yksiselitteisesti täsmälleen selittäjän keskiarvo miesten ja vastaavasti naisten osajoukossa), pätee

$$(2.16) \quad \bar{P}_m - \bar{P}_n = F_m(\bar{X}_m) - F_n(\bar{X}_n) .$$

Tässä tilanteessa otoksesta estimoitu keskipalkkaero miesten ja naisten välillä on yhtä suuri kuin tilastollisen mallin (miesten ja naisten estimoitujen työnarvostusfunktioiden) perusteella estimoitu odotettu palkkaero tyypillisen miehen ja tyypillisen naisen välillä. Muissa tapauksissa estimoitu keskipalkkaero ja odotettu palkkaero keskivertonaisen ja keskivertomiehen välillä ovat yleensä erisuuria, mutta tyypillisesti hyvin lähellä toisiaan. Kuva 2.1 havainnollistaa graafisesti keskimääräisen palkkaeron dekomponointia (vrt. Vartia ym. 1992, 85).

---

<sup>1</sup> Vastaavantyyppinen dekompositio syntyy esim. kuluttajan teoriassa, kun hyödykkeiden hintojen muutoksen seurauksena optimaalinen hyödykekori muuttuu. Tällöin siirto vanhasta optimivalinnasta uuteen voidaan ilmaista ns. tulo- ja substituutiovaikutuksen summana kahdella eri tavalla.



KUVA 2.1. Miesten ja naisten työnarvostusfunktiot

Kuvassa  $N$  on tyypillisen naisen ja  $M$  tyypillisen miehen työtehtävä. Sukupuolten välinen odotettu keskipalkkaero on kuviossa:

$$c-a = (c-d) + (d-a) = (c-b) + (b-a) .$$

Naisten keskiarvon kohdalla odotetun palkkadiskriminaation suuruus on  $(d-a)$ , ja  $(c-d)$  on työtehtävistä johtuva odotettu palkkaero miesten työnarvostusfunktiolla mitattuna. Vastavasti  $(c-b)$  on odotettu palkkadiskriminaatio miesten keskiarvon kohdalla ja  $(b-a)$  on työtehtävistä johtuva odotettu palkkaero naisten työnarvostusfunktiolla mitattuna.

### 3 TUTKIMUSAINEISTO JA MUUTTUJAT

#### 3.1 TUTKIMUSAINEISTO

Tutkimuksessa on käytetty osaa laajasta aineistosta, joka on poimittu ETLA:ssa Teollisuuden ja Työnantajain Keskusliiton palkkatilastoista. Kyseiset tilastot kattavat kaikki TT:n jäsenyrityksissä kunakin vuonna tietyinä ajankohtana työskentelevät toimihenkilöt. Vuoteen 1992 saakka otoksen poimimisajankohta on ollut elokuu ja vuonna 1993 joulukuu. Kokonaisuudessaan ETLA:n aineisto koostuu neljästätoista otoksesta ajanjaksolta 1980-1993,<sup>1</sup> jotka on poimittu kunkin vuoden TT:n aineistosta siten, että aineisto sisältää paneelia, eli samojen henkilöiden palkkoja ja muita tilastoituja ominaisuuksia on seurattu vuodesta toiseen. Samalla minkä tahansa yksittäisen vuoden otos on edustava ja sitä voidaan käyttää tutkimuksessa erillisenä poikkileikkausotoksena. Seuraavassa esitän lyhyesti miten tämä on tehty.<sup>2</sup>

Ensimmäinen otos on poimittu vuonna 1990. Palkkatilaston yritykset ovat sopimusaloittain satunnaisjärjestyksessä ja kunkin yrityksen sisällä täystyöllistetyt yli 18-vuotiaat toimihenkilöt on järjestetty kokonaispalkan mukaan. Sitten otokseen on poimittu joka 15. toimihenkilö. Näin suoritettuna systemaattisen otannan tuloksena on saatu edustava otos täysikäisistä, kokopäivätyössä olevista TT:n jäsenyritysten toimihenkilöistä. Olkoon koko palkkatilaston toimihenkilöiden määrä vuonna 1990  $N_{90}$  ja poimitun otoksen suuruus  $O_{90}$ . Otossuhteeksi saadaan  $K = O_{90} / N_{90} \approx 0.066$ . Vuoden 1991 otoksen poiminnassa palkkatilasto on jaettu kahteen osaan. Toinen osa koostuu "vanhoista toimihenkilöistä", eli kaikista niistä toimihenkilöistä, jotka ovat olleet palkkatilastossa vuonna 1990. Olkoon heidän lukumääränsä  $V_{91}$ . Loput vuoden 1991 toimihenkilöistä ovat siinä mielessä uusia, että heitä ei ollut tilastossa vuonna 1990. Uusien toimihenkilöiden lukumäärää merkitään  $U_{91}$ :lla. Vuoden 1991 otoksen poiminnassa on etsitty ne toimihenkilöt, jotka ovat mukana edellisen vuoden otoksessa.

---

<sup>1</sup>Aineistoa päivitetään jatkuvasti. Mainittu ajanjakso heijastaa tilannetta vuoden 1995 alussa.

<sup>2</sup>Tarkempi selostus tästä Asplundin, Kettusen ja Vartiainen kehittämästä poimintamenetelmästä ks. Vartiainen (1992)

Olkoon heitä  $L_{91}$  ja he tulevat automaattisesti mukaan otokseen. Vuoden 1990 otoksen edustavuus takaa, että likimain pätee yhtälö  $L_{91}/V_{91}=K$ . Seuraavaksi otosta täydennetään poimimalla uusien toimihenkilöiden osajoukosta niin paljon havaintoja (sanotaan  $T_{91}$  kappaletta), että otantasuhde pysyy muuttumattomana, eli  $O_{91}/N_{91}=K$ , jossa  $N_{91}$  on toimihenkilöiden lukumäärä vuoden 1991 palkkatilastossa. On helppo nähdä, että vuonna 1991 otoksessa toteutuu ainakin likimain ehto  $T_{91}/U_{91}=K$ . Näin ollen vanhojen ja uusien toimihenkilöiden suhteellinen osuus vuoden 1991 otoksessa on likimain sama kuin heidän osuutensa koko palkkatilastosta (likimain pätee  $L_{91}/T_{91}=V_{91}/U_{91}$ ). Tästä seuraa, että vuoden 1991 otosta voidaan pitää edustavana poikkileikkauksena vuonna 1991 TT:n palkkatilaston toimihenkilöille ja samalla otoksen vanhoista toimihenkilöistä voidaan muodostaa paneeliaineisto vuosilta 1990 ja 1991. Vastaavalla tavalla on poimittu aineisto seuraavalta vuodelta, kun nyt vanhoiksi toimihenkilöiksi lasketaan toimihenkilöt, jotka ovat palkkatilastossa vuonna 1991. Tapa, jolla poimitaan otokset vuotta 1990 edeltäviltä vuosilta, on analoginen.

Tutkielmassa on käytetty ainoastaan osaa tästä laajasta aineistosta. Tutkittaessa palkkadiskriminaatiota toimihenkilöiden keskuudessa vuonna 1993 tutkimusaineistona on luonnollisesti vuoden 1993 otos. Otoksessa on 6537 toimihenkilöä, jotka edustavat kaikkia vuonna 1993 TT:n jäsenyrityksissä työskenteleviä, kokopäivätyössä olevia täysikäisiä toimihenkilöitä, eli noin 98 000 toimihenkilöä.

Tutkielman luvussa 6 pohditaan eräiden ei-havaittavien (ei-mitattavien) henkilökohtaisten tekijöiden kuten lahjakkuuden ja ahkeruuden vaikutusta palkanmuodostukseen ja palkkadiskriminaation arviointiin ja myös yritetään huomioida analyysissä nämä tekijät. Luvun 6 tutkimusaineistona on paneeli, joka koostuu sekä vuoden 1991 että vuoden 1993 otoksessa olevista toimihenkilöistä. Tällaisia toimihenkilöitä on 5464 ja jokaisesta on kaksi havaintoa, vuodelta 1991 ja vuodelta 1993. Tarkasti ottaen he edustavat täysikäisiä, kokopäivätyössä olevia toimihenkilöitä, jotka olivat TT:n jäsenyritysten palveluksessa elokuussa 1991 ja joulukuussa 1993. On huomattava, että aineiston poimintatavan vuoksi nämä toimihenkilöt ovat olleet TT:n jäsenyritysten palveluksessa myös elokuussa vuonna 1992 (erittäin suurella todennäköisyydellä). Koska TT:n jäseniä ovat lähes kaikki teollisuusyritykset, voidaan

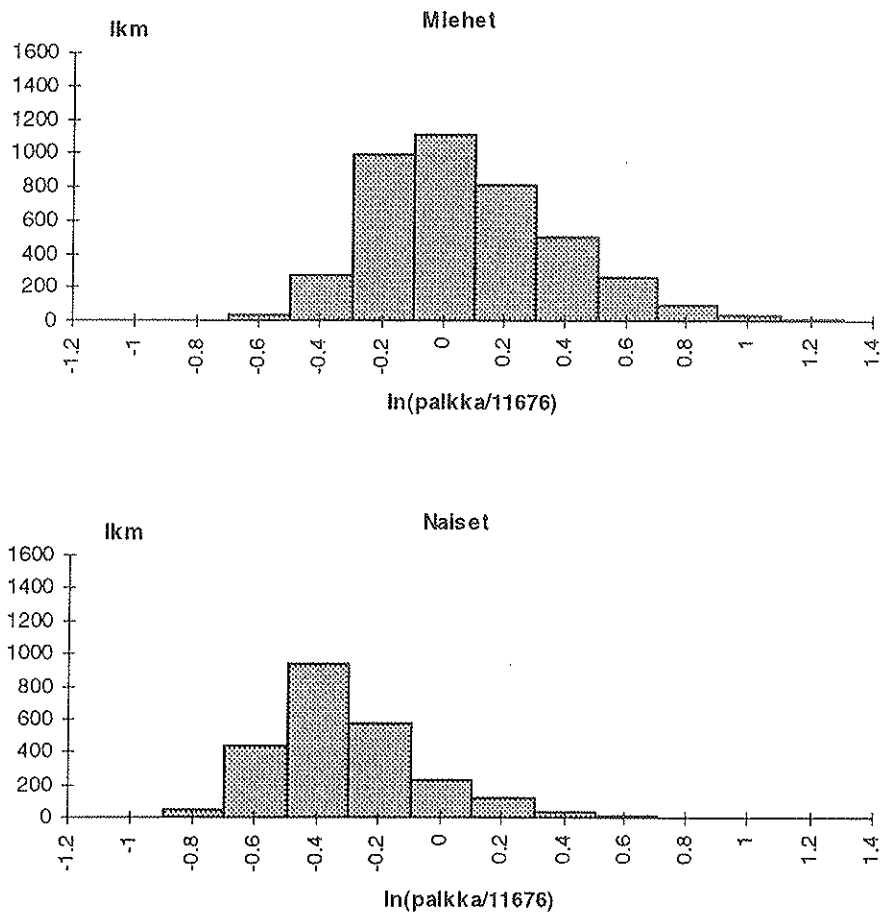
katsoa, että paneeliaineisto edustaa aikuisia Suomen teollisuudessa ajanjaksolla 1991-1993 työskennelleitä toimihenkilöitä, jotka olivat kokopäivätyössä ainakin suurimman osan ajasta.

Vuoden 1993 otosta kutsutaan yleensä jatkossa poikkileikkausaineistoksi tai poikkileikkaukseksi ja vastaavasti vuoden 1991 ja vuoden 1993 otosten yhteisistä toimihenkilöistä koostuvaa aineistoa, jossa jokaisesta toimihenkilöstä on kaksi havaintoa, sanotaan paneeliaineistoksi tai paneeliksi.

### **3.2 MUUTTUJAT**

TT:n palkkatilasto on varsin runsas. Jokaisesta henkilöstä löytyy paitsi palkkatietoja myös tietoa muodollisesta koulutuksesta, työtehtävän luonteesta ja sen vaativuudesta, toimihenkilön iästä, työkokemusvuosista, sopimusalaista, toimihenkilöryhmästä ja sukupuolesta. TT:n palkkatilastoa on laajasti selostanut Kettunen ym.(1992). Koska Vartia ym. (1992) tutkii niin ikään palkkadiskriminaatiota TT:n palkkatilastoista poimitulla otoksella, ovat tämän tutkimuksen muuttujat suurelta osin aivan samanlaisia. Useat muuttujat eivät ole sellaisenaan käyttökelpoisia tilastollisia palkkamalleja rakennettaessa, vaan niitä täytyy muokata sopivasti. Seuraavaksi esitetään tutkimuksen kannalta keskeisimmät muuttujat, jotka ovat sekä poikkileikkauksessa että paneelissa samat. Lyhyt kuvaus kaikista tutkimuksen muuttujista on liitteessä 1.

**Selitettävä muuttuja.** Tutkimuksessa selitettävänä muuttujana on toimihenkilön kunkin vuoden kokonaispalkan logaritminen poikkeama kunkin vuoden toimihenkilöiden keskipalkasta, joka on laskettu koko TT:n tilastosta. Toimihenkilön kokonaispalkkaan sisältyvät kiinteän kuukausipalkan lisäksi vuorotyölisät, luontoisetujen raha-arvo ja tulos- ja suoritussidonnaiset erät. Kokonaispalkan ja keskipalkan suhde kuvaa toimihenkilön suhteellista palkkaetua, joka on sama riippumatta siitä, ovatko kyseessä reaali- vai nominaalipalkat. Tästä on hyötyä erityisesti, kun tutkimusaineisto on paneeli. Kuvassa 3.1 esitetään selitettävän muuttujan jakauma poikkileikkausotoksessa sukupuolen mukaan.



KUVA 3.1. Selitettävän muuttujan jakauma poikkileikkausaineistossa sukupuolen mukaan

Naiset ovat selvästi huonommin palkattuja kuin miehet. Selitettävän muuttujan keskiarvo naisten osajoukossa on noin -0.31 ja miesten osajoukossa noin 0.07. Miesten ja naisten kokonaispalkkaero on keskimäärin 38.6 log-%<sup>1</sup>, eli naisten palkka on keskimäärin noin 32 % pienempi kuin miesten palkka. Paneeliaineistossa näkyy sama tilanne. Paneelissa olevat toimihenkilöiden palkat vuonna 1993 ovat keskimäärin hieman korkeampia kuin poikkileikkausaineiston toimihenkilöiden (ero on todella pieni) ja miesten palkkojen etumatka naisten palkkoihin verrattuna säilyy paneelissa ennallaan.

<sup>1</sup> Kyseessä on miesten ja naisten geometrinen keskipalkkojen suhteellinen ero log-prosentteissa ilmaistuna.



**Toimihenkilöryhmät.** TT luokittelee toimihenkilöt kolmeen suureen ryhmään. Nämä ovat konttoritoimihenkilöt (vuodesta 1993 teollisuustoimihenkilöt), tekniset toimihenkilöt ja ylempät toimihenkilöt. Jako perustuu siihen, että eri ryhmissä olevat henkilöt kuuluvat eri työehtosopimukseen, joten on luontevaa ajatella, että palkanmuodostus saattaisi ainakin osittain poiketa toimihenkilöryhmittäin. Taulukko 3.1 esittää poikkileikkaus- ja paneeliaineiston toimihenkilöiden jakautumisen toimihenkilöryhmittäin.

	Lukumäärä			Suhteellinen osuus, %	
	miehet	naiset	yhteensä	miehet	naiset
<b>V. 1993 poikkileikkaus</b>					
Konttoritoimihenkilöt	359	1666	2025	9	69
Tekniset toimihenkilöt	1912	396	2308	46	16
Ylemmät toimihenkilöt	1857	347	2204	45	14
Yhteensä	4128	2409	6537	100	100
<b>Paneeliaineisto</b>					
Konttoritoimihenkilöt	292	1375	1667	8	69
Tekniset toimihenkilöt	1631	334	1965	47	17
Ylemmät toimihenkilöt	1556	276	1832	45	14
Yhteensä	3479	1985	5464	100	100

**TAULUKKO 3.1.** Toimihenkilöt toimihenkilöryhmittäin

Naistoimihenkilöt ovat pääasiassa konttoritoimihenkilöitä. Miehet jakautuvat tasaisesti tekniisiin ja ylempiin toimihenkilöihin, vain vajaat 10 % miehistä on konttoritoimihenkilöitä. Toimihenkilöiden jakauma toimihenkilöryhmittäin on poikkileikkauksessa ja paneelissa sama.

**Toimiperheet ja nimikkeet.** TT:n palkkatilastoissa toimihenkilöt on jaettu 29 toimiperheeseen. Samassa toimiperheessä olevien toimihenkilöiden työtehtävät kuuluvat laajasti ottaen samaan työtehtäväpiiriin. Kullakin toimiperheellä on kaksinumeroinen koodi. Esimerkiksi

toimiperhe 41 tarkoittaa myyntityötä. Toimiperheen sisällä toimihenkilöt on jaettu eri nimikkeisiin, joilla on kolminumeroinen koodi. Kolminumeroisen koodin ensimmäiset kaksi numeroa osoittavat, mihin toimiperheeseen toimihenkilö kuuluu. Kaikkiaan TT:n tilastoissa on 79 nimikettä.

Esimerkkinä myyntityön toimiperheen nimikkeet ja niiden koodit esitetään taulukossa 3.2.

Kolminumeroinen nimike	Koodi
Myynnin johtotyö	411
Erikoismyyntityö ja välillinen myyntityö	412
Vientimyyntityö	413
Myyntityö	414
Toimistomyyntityö ja avustava myyntityö	415
Tuote-esittelytyö ja asiakaspalvelutyö	416

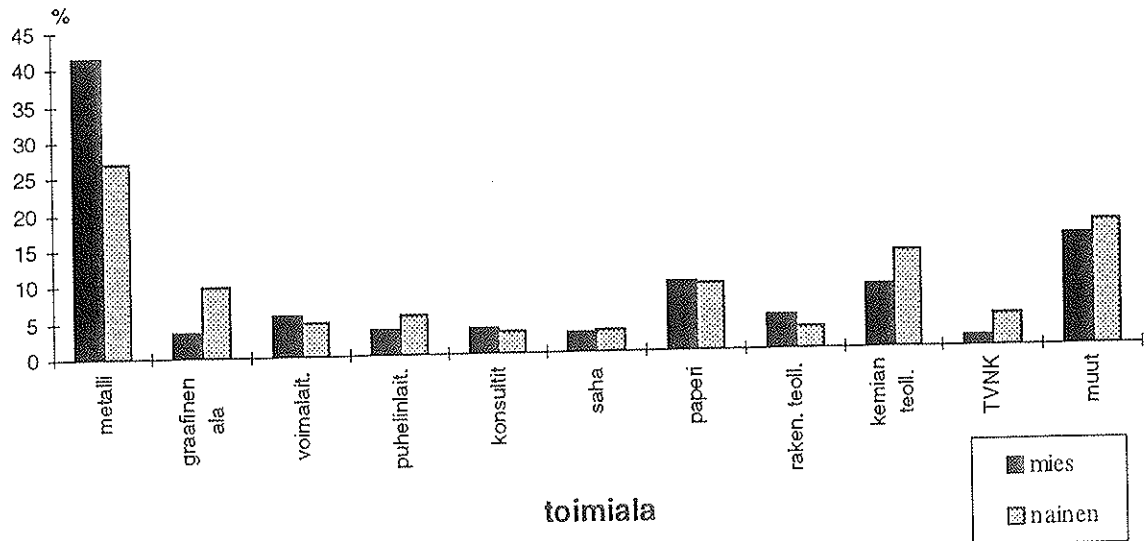
**TAULUKKO 3.2.** *Kolminumeroiset nimikkeet toimiperheessä 41 (myyntityö)*

Kolminumeroinen nimike kuvaa karkeasti toimihenkilön työtehtävää, jatkossa siihen viitataan nimellä kolminumeroinen työtehtävä. Palkanmuodostuksen kannalta työtehtävä on keskeinen, joten palkkayhtälöissä on aivan välttämätöntä huomioida, että palkkataso eri nimikkeissä on erilainen. Koska nimikkeitä on paljon, analyysi vaikeutuu huomattavasti, jos jokaiselle nimikkeelle luodaan indikaattorimuuttuja. Tästä syystä on rakennettu muuttuja **lpoikkp**, joka ilmoittaa tiettyyn kolminumeroiseen työtehtävään kuuluvien toimihenkilöiden keskipalkan logaritmisestä poikkeaman kaikkien toimihenkilöiden keskipalkasta. Keskipalkat on laskettu vuoden 1993 koko palkkatilastosta. Paneeliaineistossa tarvittavat keskipalkat ja niiden logaritmiset poikkeamat lasketaan vuosille 1991 ja 1993 erikseen. Muuttuja **lpoikkp** kuvastaa eri tehtävien suhteellista arvostusta työmarkkinoilla. Jos hyväksytään ajatus, että vaativammat tehtävät ovat myös paremmin palkattuja, kuvastaa **lpoikkp** myös työtehtävien suhteellista vaativuustasoa.

**Toimialat.** Aineiston toimihenkilöt edustavat suurta joukkoa sopimusaloja, joista jotkut ovat todella suuria, kuten metalliala, johon kuuluu runsas kolmasosa otosten toimihenkilöistä. Toisaalta tiiliteollisuustoimihenkilöitä on vain kaksi. Sen takia muodostettiin uusi

muuttuja ULI, jossa eräät sopimusalat yhdistettiin. ULI jakaa toimihenkilöt yhteentoista ryhmään siten, että samaan ryhmään kuuluvien toimihenkilöiden voidaan katsoa työskentelevän samalla toimialalla. Toimialojen muodostaminen esitetään liitteessä 2.

Esitetty toimialoittainen jako on harkinnanvarainen, mutta TT:n aineisto ei anna tässä suhteessa parempia mahdollisuuksia eikä toimialojen määrittelemisen ole edes periaatteessa yksikäsitteistä. Kuva 3.2 esittää toimihenkilöiden jakauman toimialoittain vuoden 1993 poikkileikkausaineistossa. Paneeliaineistossa toimihenkilöt jakautuvat sopimusaloihin täysin samalla tavalla.



KUVA 3.2. Poikkileikkausaineiston toimihenkilöiden jakauma toimialoittain

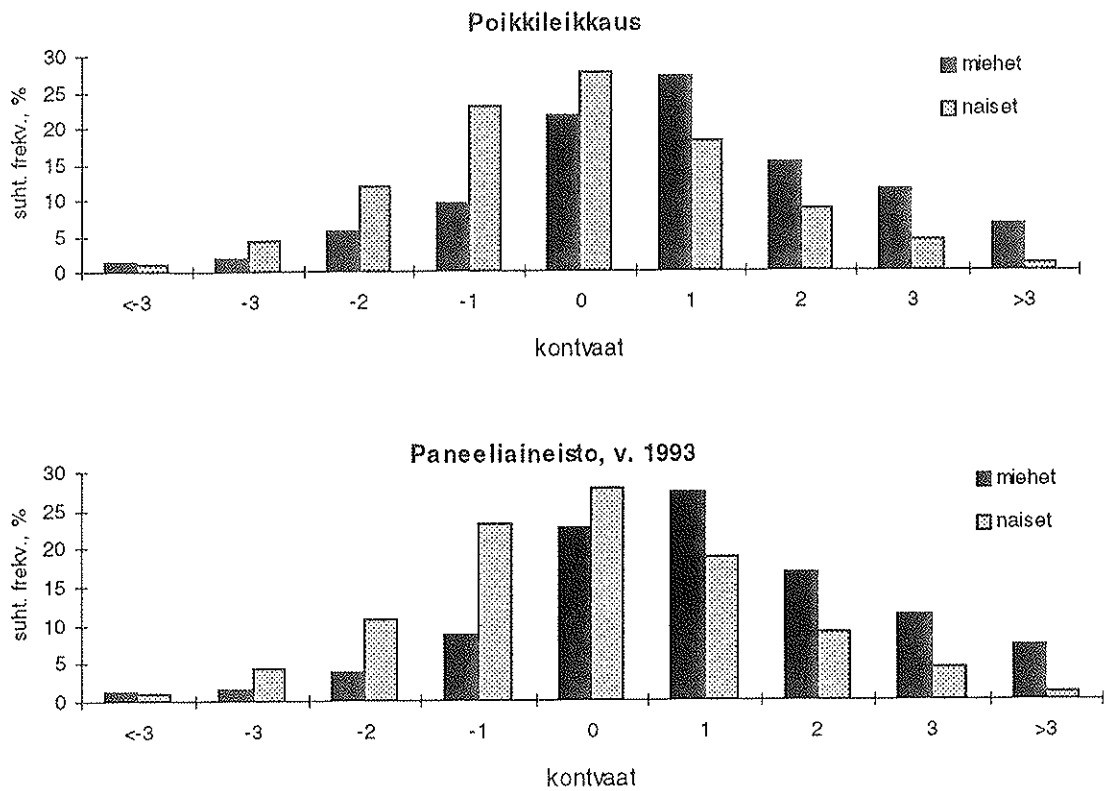
Toimihenkilöt ovat voimakkaasti keskittyneet metallialalle, miehet voimakkaammin kuin naiset. Muita suhteellisen suuria aloja ovat paperiteollisuus ja kemianteollisuus. Toimialoittaisen jaon perusteella on muodostettu muuttuja **lulikps** - toimialan toimihenkilöiden keskipalkan logaritminen poikkeama kaikkien toimihenkilöiden keskipalkasta. **Lulikps** kuvaa toimialojen välisiä suhteellisia palkkaeroja. Teknisesti muuttuja on muodostettu täysin vastaavalla tavalla kuin edellisessä kohdassa muuttuja **lpoikkp**, nyt kuitenkin toimialoittaiset keskipalkat on laskettu otoksesta.

**Työn vaativuutta kuvaavat muuttujat.** Konttoritoimihenkilöiden osalta on TT:n palkkatileastoista ilmoitettu toimen vaativuustaso yleissopimuksen (STK-STL) määrittelemän 12-asteisen luokituksen mukaan. Teknisten toimihenkilöiden osalta TT:n tilasto ilmoittaa palkkaryhmän. Palkkaryhmiä on runkosopimuksen (STK-STTK-T) mukaan kuusi. Ylempiä toimihenkilöitä ei ole luokiteltu vastaavan toimen (työtehtävän) vaativuustasoa kuvaavan muuttujan mukaan. Alkuperäisen tilaston palkkaryhmien ja vaativuustasojen esitysmuodot eivät sellaisenaan sovi tilastollisen analyysin muuttujiksi, joten ne on esitettävä palkkamallien rakentamista varten soveltuvassa muodossa. Alkuperäiset vaativuustasojen koodit ja niiden uudet, analyysia varten luodut numeeriset arvot esitetään liitteessä 2. Numeerisessa muodossa esitettynä konttoritoimihenkilöiden vaativuustasot muodostavat muuttujan **kontvaat**, jonka arvot konttoritoimihenkilöiden kohdalla ovat kokonaislukuja väliltä  $\{-6, 5\}$ , matalimman vaativuustason ollessa -6 ja korkeimman 5. Vastaavasti numeerisesti esitetyt teknisten toimihenkilöiden palkkaryhmät muodostavat muuttujan **teknvaat**, jonka arvot ovat teknisten toimihenkilöiden kohdalla kokonaislukuja väliltä  $\{-3, 2\}$ , siten että isompi luku vastaa korkeampaa palkkaluokkaa, eli teknisen työtehtävän vaativuustasoa.

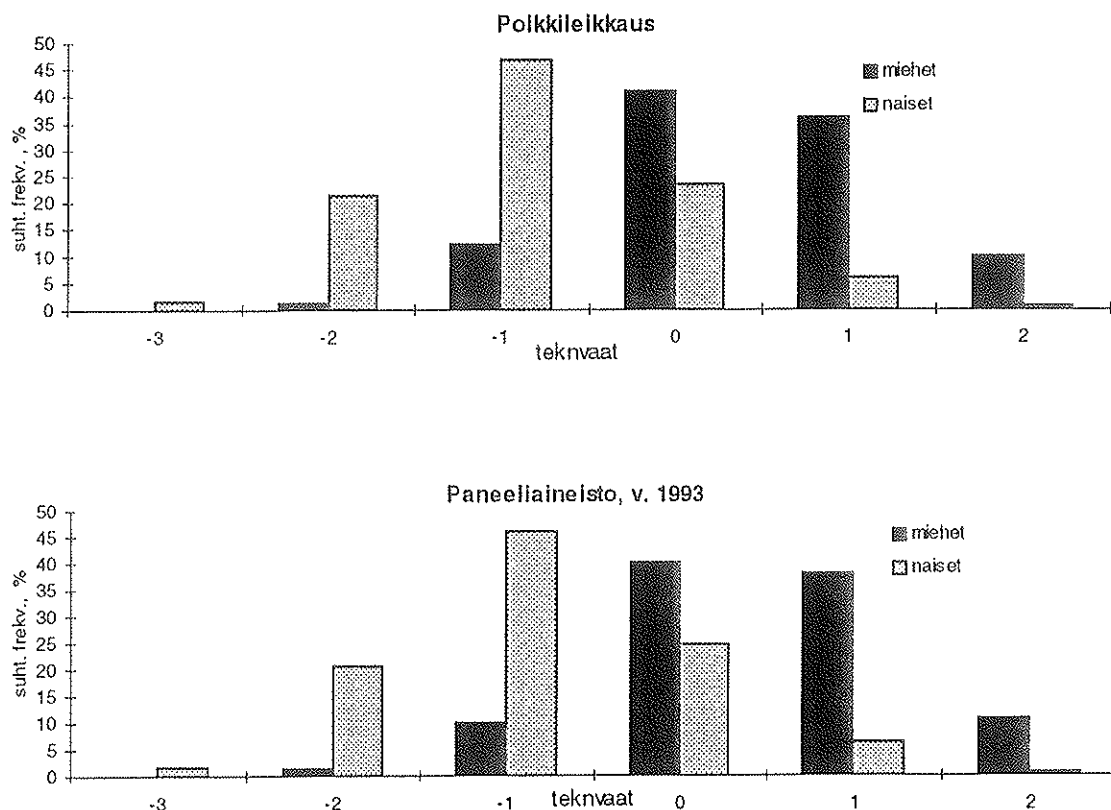
Analyysin kannalta on ongelmallista, että eri toimihenkilöryhmissä työtehtävän vaativuustasoa kuvaavat muuttujat ovat erilaisia. Vaativuustason eri versioissa on siis runsaasti puuttuvia tietoja. Vartia ym. (1992) ratkaisee ongelman antamalla muuttujalle **kontvaat** teknisten ja ylempien toimihenkilöiden kohdalla arvoksi sen otoskeskiarvon konttoritoimihenkilöiden osajoukossa. Melkein samaan johtaisi menetelmä, jos puuttuvat tiedot korvattaisiin indikaattorimuuttujilla. Tämä "pidentäisi" tarpeettomasti regressiomalleja ja vaikeuttaisi tulkintoja. Vastaavasti muuttujan **teknvaat** arvo konttori- ja ylempien toimihenkilöiden kohdalla on sen otoskeskiarvo teknisten toimihenkilöiden osajoukossa.

Paneeliaineistossa kunkin vuoden puuttuvat tiedot on korvattu muuttujien **kontvaat** ja **teknvaat** keskiarvolla kultakin vuodelta (keskiarvo on laskettu aina vastaavan vuoden poikkileikkausaineiston perusteella), eli esimerkiksi ylempien toimihenkilöiden kohdalla muuttujat **teknvaat** ja **kontvaat** saavat hieman eri arvon vuonna 1991 kuin vuonna 1993

(ero on varsin pieni). Näin on menetelty, koska vaativuustasoa kuvaavien muuttujien jakauma ei ole sama eri ajankohtina. Matalan vaativuustason työtehtävissä olevien toimihenkilöiden osuus vähenee ajan kuluessa (ks. Asplund 1993b, 53-54).



KUVA 3.3a. Muuttujan kontvaat jakauma konttoritoimihenkilöiden osajoukossa



KUVA 3.3b. Muuttujan *teknyaat* jakauma teknisten toimihenkilöiden osajoukossa

Kuva 3.3a esittää muuttujan *kontvaat* jakauman konttoritoimihenkilöiden osajoukossa poikkileikkaus- ja paneeliaineistossa (vuonna 1993) sukupuolen mukaan. Kuva 3.3b on vastaavanlainen esitys muuttujan *teknyaat* jakaumasta teknisten toimihenkilöiden osajoukossa. Poikkileikkausaineistossa muuttujan *kontvaat* keskiarvo on miesten kohdalla 0.7 ja naisten kohdalla -0.11. Vastaavasti muuttujan *teknyaat* keskiarvo on miesten kohdalla 0.41, kun naisten kohdalla keskiarvo on -0.1. Naiset siis sijoittuvat sekä teknisten että konttoritoimihenkilöiden ryhmässä vähemmän vaativiin tehtäviin kuin miehet. Tilanne on sama paneeliaineistossa.

Edellä esitetyt muuttujat *lpoikkp*, *kontvaat* ja *teknyaat* ovat palkanmuodostuksen kannalta keskeisimmät. Tiettyyn työtehtävään ja palkkaryhmään kuulumisen voidaan katsoa määrittävän peruspalkan. Muuttujat, kuten ikä (työkokemus) ja koulutus, selittävät työtehtävien

sisäisiä palkkaeroja, jotka yleensä ovat huomattavasti pienempiä kuin työtehtävien väliset palkkaerot (Vartia 1992, 39).

**Muodollinen koulutus.** Koulutus on keskeisin inhimillistä pääomaa kuvaava tekijä. TT:n palkkatilastot ilmoittavat toimihenkilön muodollisen koulutuksen kaksinumeroisena koodina. Ensimmäinen numero ilmoittaa koulutusalan ja toinen numero koulutustason. Koulutusalan vaikutusta palkanmuodostukseen ei ole tutkittu tarkasti tässä tutkielmassa, koska alustavien tarkastelujen ja aiempien tutkimusten perusteella voidaan katsoa, että koulutusala vaikuttaa pääasiassa sijoittumiseen tietyn tyyppiseen työtehtävään. Taulukko 3.3 esittää koulutustason seitsemänasteisen luokituksen ja toimihenkilöiden jakautumisen koulutustason mukaan.

Koulutustaso	Luokitus	Poikkileikkaus		Paneeli	
		Toimihenkilöiden suhteellinen osuus %		Toimihenkilöiden suhteellinen osuus %	
		Mies	Nainen	Mies	Nainen
Kansa- ja kansalaiskoulu	1	6.10	11.08	6.21	11.89
Keski- tai peruskoulu, alempi keskiaste	2	15.12	30.14	15.12	32.29
Ylioppilastutkinto ja opistoaste	3	41.09	45.04	41.56	42.97
Alempi korkeakoulututkinto	4	22.60	7.47	22.62	7.46
Ylempi korkeakoulututkinto (maisteri)	5	14.12	5.89	13.54	5.04
Lisensiaatin tutkinto	6	0.68	0.21	0.63	0.15
Tohtorin tutkinto	7	0.29	0.17	0.32	0.20
<b>Keskimääräinen koulutustaso</b>		<b>3.26</b>	<b>2.68</b>	<b>3.25</b>	<b>2.63</b>

**TAULUKKO 3.3.** *Koulutustason luokitus ja aineistojen toimihenkilöiden suhteelliset osuudet eri koulutustasoilla*

Taulukosta ilmenee, että paneeli- ja poikkileikkausaineiston toimihenkilöt jakautuvat vuonna 1993 samalla tavalla koulutustason mukaan. Miehet ovat keskimäärin naisia paremmin koulutettuja. Keskimäärin miesten muodollinen koulutus on jonkin verran ylioppilastutkintotasoista koulutusta korkeampi, kun taas naisten muodollinen koulutus on keskimäärin alemman keskiasteen ja ylioppilastutkinnon välillä.

**Ikä ja työkokemus.** Muodollisen koulutuksen ohella työkokemus on toinen keskeinen inhimillistä pääomaa kuvaava tekijä, koska työkokemus kasvattaa henkilön osaamista.

Palkkamalleissa henkilön ikää käytetään yleensä työkokemuksen sijasta. Tausta-ajatuksena on, että iäkkäämmät henkilöt ovat yleensä olleet pitempään työelämässä. Palkkadiskriminaatiotutkimuksessa iän käyttö selittäjänä palkkamalleissa saattaa olla ongelmallinen, jos naisten työura on esim. perhesyistä katkonaisempi kuin miesten. Tällöin samanikäiset miehet ja naiset eivät ole keskimäärin yhtä kokeneita työntekijöitä, mikä voi johtaa jossain määrin virheelliseen palkkadiskriminaation arviointiin.<sup>1</sup> TT:n aineistossa on tietoa työkokemuksesta ainoastaan teknisten ja konttoritoimihenkilöiden osalta. Ylemmille toimihenkilöille joudutaan laskemaan ns. potentiaalinen työkokemus, joka saadaan kaavasta<sup>2</sup>

(3.1) potent. työkokemus = ikä - koulutusvuodet - koulunkäynnin aloitusikä (Suomessa 7 v.)

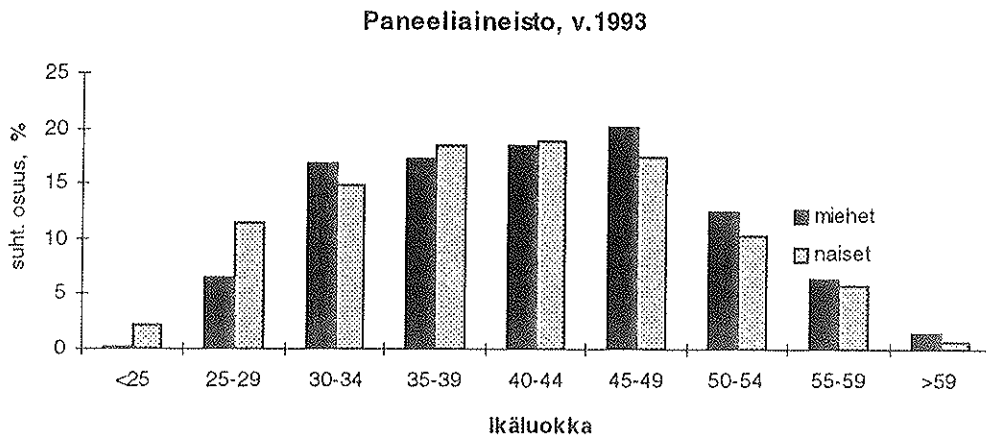
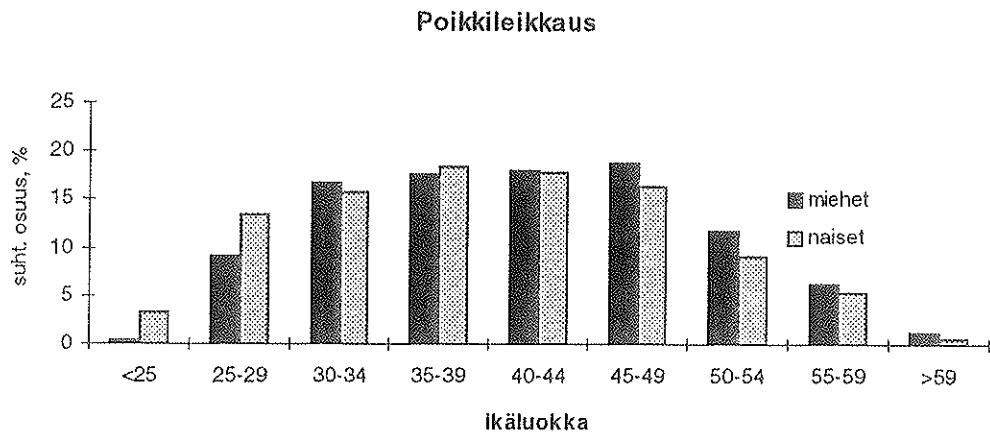
Potentiaalinen työkokemus ei ole tutkimuksen tarkoituksen kannalta välttämättä parempi selittäjä kuin tavallinen ikä. Mutta myös konttori- ja teknisten toimihenkilöiden osalta TT:n tilastossa ilmoitettuihin työkokemusvuosiin täytyy suhtautua jossain määrin varauksellisesti, koska ne ovat itse asiassa palkkauksen perusteena olevia kokemusvuosia (TT:n jäsentiedote 1991, 4), joiden laskeminen on määritelty työehtosopimuksissa. Monessa tapauksessa ne eivät ole yhtä kuin todellinen työssäoloaika. Tämän tutkimuksen kannalta ehkä oleellisinta on se, että äitiysloma ja vanhempainloma, joiden yhteinen kesto on noin 10 kuukautta lasta kohti, lasketaan työkokemukseksi (tavallisessa kielenkäytössä äitiyslomasta puhuttaessa tarkoitetaan usein sekä äitiyslomaa että vanhempainlomaa). Tämän vastapainoksi mainittakoon, että miesten kohdalla asepalvelun ei katsota keskeyttävän työsuhdetta, joten asepalvelusaika on mukana palkkaperusteena olevissa työkokemusvuosissa sellaisten miesten kohdalla, jotka ovat olleet samassa työsuhteessa ennen armeijaan menoa. Tarkempi tieto palkkauksen perusteena olevien työkokemusvuosien määräytymisestä löytyy teknisten toimihenkilöiden osalta STK-STTK-T:n välisen runkosopimuksen liitteenä olevasta palkkasopimuk-

<sup>1</sup> Esiintyy tendenssi yliarvioida todellista palkkadiskriminaatiota (ks. Oaxaca, 1973, 697).

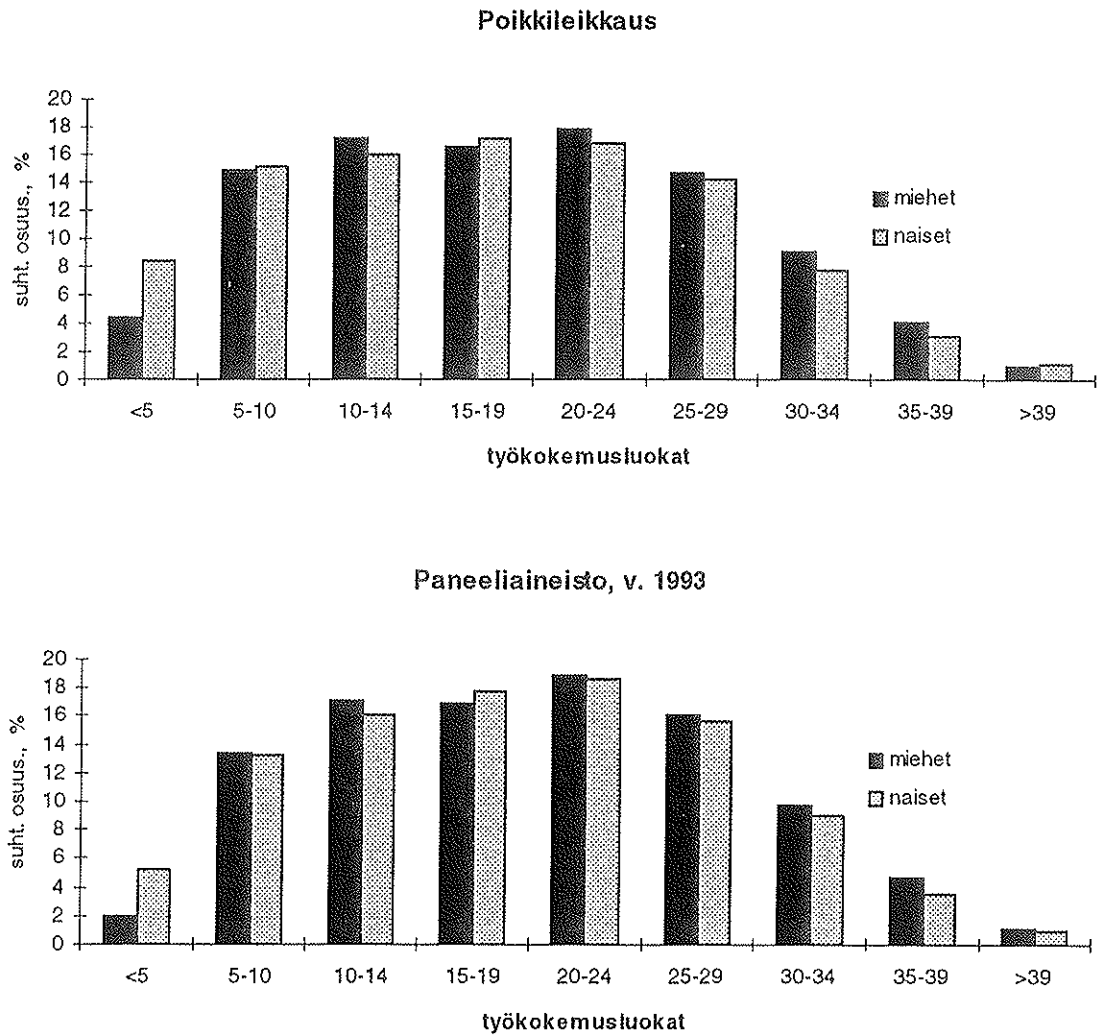
<sup>2</sup> Käytetään Asplundin (1993b) esittämää koulutusavainta, jolla muunnetaan suoritettu tutkinto koulutusvuosiksi seuraavasti: kansa- ja kansalaiskoulu = 7 v., keski- ja peruskoulu = 9 v., alempi keskiaste = 11 v., ylioppilastutkinto = 12 v., alempi korkeakoulututkinto = 14 v., ylempi korkeakoulututkinto = 16 v., lisensiaattitutkinto = 18 v., tohtorin tutkinto = 20 v.



sen soveltamisohjeesta ja STK-STL:n välisen yleissopimuksen liitteenä olevasta palkkausjärjestelmästä. Seuraavaksi esitetään toimihenkilöiden ikäjakauma (Kuva 3.4) ja työkokemusjakauma (Kuva 3.5).



KUVA 3.4. Toimihenkilöiden ikäjakauma



KUVA 3.5. Toimihenkilöiden työkokemusjakauma

Keskimääräinen ikä v. 1993 on miesten kohdalla 41.3 vuotta poikkileikkauksessa ja 42 vuotta paneeliaineistossa. Naisten kohdalla vastaavat luvut ovat 39.5 ja 40.4 vuotta. Naiset ovat keskimäärin miehiä kaksi vuotta nuorempia ja paneelissa olevat toimihenkilöt ovat vuoden verran poikkileikkauksessa olevia toimihenkilöitä vanhempia. Miehillä on poikkileikkausaineistossa keskimäärin 18.8 työkokemusvuotta ja paneelissa 19.7 työkokemusvuotta. Naisten kohdalla luvut ovat 17.8 ja 18.9 työkokemusvuotta. Toisin sanoen ikäero miesten ja naisten välillä näyttää olevan suurempi kuin sukupuolten välinen työkokemusero, mikä on varsin yllättävää ottaen huomioon, että on vahvat syyt ajatella naisten työuran ole-

van keskimäärin miesten työuraa katkonaisempi.<sup>1</sup> Tilanne ei muutu, jos tarkasteluista suljetaan pois ylemmät toimihenkilöt, joiden kohdalla työkokemus on täysin laskennallinen. Syynä on luultavasti se, että ilmoitetut palkkaperusteena olevat työkokemusvuodet eivät mittaa riittävän hyvin todellisia työkokemusvuosia. Ainakin on selkeitä syitä ajatella näin, kuten edellä esitetystä ilmenee. Joka tapauksessa on mielenkiintoista tutkia, muuttuuko käsitys palkkadiskriminaatiosta, jos iän sijasta palkkamalleissa käytetään selittäjänä aineiston tarjoamaa työkokemusmuuttujaa. Siksi on konstruoitu vaihtoehtoiset palkkamallit, joissa selittäjänä on joko ikä tai työkokemus.

### **3.3 YHTEENVETO**

Kaikki jatkoanalyysin kannalta tärkeimmät muuttujat, jotka esitettiin tässä luvussa, on pyritty muodostamaan niin, että ne olisivat mahdollisimman samanlaisia kuin Vartian ym. (1992) tutkimuksessa. Näin saadaan vertailukelpoisia tuloksia. Keskeisten muuttujien jakaumaa tarkasteltiin lyhyesti sekä poikkileikkausaineistossa että paneelissa. Poikkileikkausotoksessa olevista toimihenkilöistä on paneelissa 83.6 %. Yleisesti ottaen tarkastellut muuttujat eivät jakaudu eri tavalla, kun analysoidaan koko vuoden 1993 otosta tai ainoastaan paneelissa olevia henkilöitä. Vain iässä ja työkokemuksessa havaittiin jossain määrin isompia eroja. Paneelissa olevat toimihenkilöt ovat keskimäärin vuoden vanhempia ja kokeneempia verrattuna kaikkiin poikkileikkausotoksen toimihenkilöihin. Erot olisivat tietenkin suurempia, jos verrattaisiin paneeliaineiston toimihenkilöitä niihin toimihenkilöihin, jotka eivät ole paneelissa. Sen sijaan tarkasteltujen muuttujien jakaumassa oli selkeitä eroja sukupuolten välillä. Naiset ovat keskimäärin miehiä nuorempia ja heillä on myös vähemmän työkokemusta. Miehet ovat naisia paremmin koulutettuja ja sijoittuvat vaativampiin (ja paremmin palkattuihin) tehtäviin.

---

<sup>1</sup> On odotettavissa, että jos naisten työura on katkonaisempi, ero miesten ja naisten keskimääräisissä työkokemusvuosissa olisi keskimääräistä ikäeroa suurempi, olettaen että molemmat aloittavat työelämässä samanikäisinä.

## 4 PALKKADISKRIMINAATIOFUNKTION ESTIMOINTI

Palkkadiskriminaation empiirisen arvioinnin metodologia esitettiin luvussa 2. Tehtävänä on estimoida laaja palkkayhtälö, joka kuvaa realistisesti toimihenkilöiden palkanmuodostusta ja erityisesti palkanmuodostuksessa esiintyviä eroja sukupuolten välillä. Estimoidun palkkayhtälön perusteella konstruoidaan palkkadiskriminaatiofunktio, joka sitten ilmoittaa odotetun palkkadiskriminaation suuruuden tietyssä pisteessä palkkamallin selittäjien avaruudessa. Palkkamallin selittäjien on oltava mahdollisimman selkeästi tulkittavissa ja niiden vaikutus palkanmuodostukseen on oltava tilastollisesti merkitsevää. Tämä ns. varovaisuusperiaate takaa muun muassa sen, ettei miesten ja naisten estimoiduissa palkkayhtälöissä esiinny eroja, joita ei todellisuudessa ole olemassa.

Palkkayhtälön rakentaminen on työläs tehtävä tilanteessa, jossa aineisto on varsin heterogeeninen ja on käytettävissä paljon selitettäviä muuttujia, jotka kaikki potentiaalisesti vaikuttavat palkanmuodostukseen. Sitä ei voi suorittaa mekaanisesti, yksiselitteisiä ohjeita noudattaen. Tässä analyysin pääperiaatteena on ollut lähteä laajan mallin estimoinnista, jossa on mukana selittäjinä mm. kaikki naisindikaattorin ja muiden selittäjien interaktiotermit. Testausmenetelmiä (F- ja t-testit) käyttäen yksinkertaistetaan mallia asteittain siten, että lopullisessa palkkamallissa kaikki selittäjät ovat tilastollisesti merkitseviä ja tulkinnallisesti selkeitä. Menetelmää ei ole kuitenkaan noudatettu aivan kirjaimellisesti, vaan kaikkien toimihenkilöiden yhteistä palkkamallia rakennettaessa on käytetty apuna myös toimihenkilöryhmittäin erikseen estimoituja palkkayhtälöitä ja tietoa aikaisemmista tutkimuksista, joissa käytetään otoksia TT:n palkkatilastoista (Asplund 1993b ja Vartia ym. 1992). Testauksen tarkoituksena on ollut pääasiassa selvittää, poikkeako ja millä tavalla, keskeisten selittäjien (koulutuksen, iän, työtehtävän, työn vaativuuden) vaikutus palkkaan toimihenkilöryhmien, ja ennen kaikkea sukupuolten välillä - tämä on empiirinen kysymys. Kun tietty laaja malli yksinkertaistetaan vaiheittain, lopullisen mallin muotoon saattaa vaikuttaa järjestys, jossa hypoteeseja testataan. Siksi on kokeiltu myös erilaisia vaihtoehtoisia tapoja rajoittaa lähtökohtaiset mallit. Lopullisen mallin valintaan vaikuttivat mallin selitys-

aste (keskivirheen hajonta) ja yksittäisten selittäjien selkeä tulkinta<sup>1</sup> ja tilastollinen merkitsevyys. Estimointimenetelmänä on ollut tavallinen pienimmän neliösumman menetelmä, mutta testauksessa on käytetty heteroskedastisuuden suhteen Whiten (1980) esittämiä korjattuja kovarianssimatriiseja. Regressioajot on suoritettu LIMDEP 6.0 ohjelmistolla. Koska mallin valinta oli monivaiheinen ja on kokeiltu erilaisia vaihtoehtoisia spesifikaatioita, lopullisessa mallissa on mukana vain selittäjiä, joiden vaikutus palkanmuodostukseen on tilastollisesti merkitsevä alle 1 % virhetasolla, mikä on tavanomaista tiukempi kriteeri.

Jo lähtökohtaisesti on spesifioitu kahta eri mallityyppiä. Toisessa mallityypissä selittäjänä on henkilön ikä ja toisessa palkkaperusteena oleva työkokemus (ylempien toimihenkilöiden osalta potentiaalinen työkokemus). Vastaavasti lopullisia mallejakin on kaksi. Koska palkkadiskriminaation arvioinnin kannalta tulokset ovat lähes identtisiä, riippumatta siitä kumpaa malleista käytetään analyysissä, tässä ja seuraavassa luvussa esitetään vain tulokset, jotka perustuvat malliin, jossa selittäjänä on ikä. Näin menetellään, koska tämä malli kuvaa selityksasteen ja virrehajonnan kannalta hieman paremmin toimihenkilöiden palkanmuodostusta. Ero toiseen malliin on toki hyvin pieni. Myös Vartia ym. (1992) käyttää selittäjänä ikää. Malli, jossa selittäjänä on työkokemus, ja siihen perustuvan palkkadiskriminaatioanalyysin keskeiset tulokset esitetään liitteessä 3.

Taulukossa 4.1 esitetään, mitä toimihenkilön ominaisuuksia ja työtehtävän ominaispiirteitä on tutkittu, millä muuttujilla ja muuttujien interaktioilla, rakennettaessa palkkamalli, jossa selittäjänä on henkilön ikä. Lopullisesti valitun mallin PM estimointitulokset esitetään taulukossa 4.2.

---

<sup>1</sup>Selittäjien tulkinta ei ole tässä tutkimuksessa ongelma - kaikki analyysissä käytetyt selittäjät ovat varsin luontevia palkanmuodostuksen kannalta.

**TAULUKKO 4.1.** *Palkkamallia rakennettaessa tutkitut palkanmuodostukseen (potentiaalisesti) vaikuttavat tekijät ja niitä kuvaavat muuttujat*

Toimihenkilön ominaisuus, työtehtävän ominaispiirre tai luokitus	Muuttuja
toimihenkilöryhmä (ylemmät, konttori- ja tekniset toimihenkilöt)	ylempien toimihenkilöiden indikaattori, konttoritoimihenkilöiden indikaattori
muodollinen koulutustaso	koulutustaso, koulutustason neliöjuuri, akateemisen koulutuksen indikaattori, edellisten interaktiot ylempien ja konttoritoimihenkilöiden indikaattorien kanssa, tutkinnon suoritusvuosi
ikä	ikä, iän neliöjuuri ja niiden interaktiot toimihenkilöryhmän indikaattorien kanssa
kaksinumeroinen toimiperhe	toimiperheiden indikaattorimuuttujat
kolminumeroinen työtehtävä	lpoikkp, sen interaktiot toimihenkilöryhmän indikaattorien kanssa
työtehtävän vaativuustaso	teknvaat, kontvaat
sopimusala (toimiala)	sopimusalaindikaattorit
sukupuoli	naisindikaattori
muut	kunnan kalleusluokka, viikkotyötunnit
lisäksi	kaikkien yllä olevien muuttujien interaktiot naisindikaattorin kanssa

TAULUKKO 4.2. Lopullisen palkkamallin PM estimointitulokset

Muuttuja	Kerroin	Keskivirhe	p-arvo
vakio	-0.482		
kunnan kalleusluokka	-0.046	0.004	0.00000
ikä	-0.025	0.004	0.00000
kontvaat, kont. tehtävien vaativuustaso	0.057	0.003	0.00000
teknvaat, tekn. tehtävien vaativuustaso	0.106	0.004	0.00000
koulutustaso	0.089	0.019	0.00001
tutkinnon suoritusvuosi	-0.001	0.000	0.00023
työtunnit/viikko	-0.018	0.002	0.00000
lpoikkp, työtehtävän keskipalkan ja yleiskeskipalkan log. ero	0.118	0.026	0.00000
yl31, ylempien toimihenkilöiden indikaattori	-2.381	0.312	0.00000
ko21, konttoritoimihenkilöiden indikaattori	-0.107	0.009	0.00000
ikä*yl31	-0.046	0.008	0.00000
sqrt(koulutustaso)	-0.195	0.059	0.00125
sqrt(ikä)	0.356	0.050	0.00000
sqrt(ikä)*yl31	0.698	0.099	0.00000
akat. koulutus*yl31	0.100	0.011	0.00000
lpoikkp*yl31	0.350	0.030	0.00000
lpoikkp*ko21	0.215	0.035	0.00000
toimiperhe12	-0.040	0.006	0.00000
toimiperhe14	-0.066	0.010	0.00000
toimiperhe21	0.027	0.006	0.00007
toimiperhe41	0.049	0.008	0.00002
graafinen ala	0.159	0.013	0.00000
voimalaitokset	0.108	0.011	0.00000
puhelinlaitokset	0.054	0.010	0.00000
konsultit	0.039	0.011	0.00018
sahateollisuus	-0.055	0.012	0.00000
paperiteollisuus	0.130	0.009	0.00000
rakennusteollisuus	0.074	0.010	0.00000
kemianteollisuus	0.083	0.009	0.00000
muut alat	0.059	0.007	0.00000
N, naisindikaattori	-0.578	0.125	0.00000
N*yl31	0.236	0.044	0.00001
N*yl31*ikä	-0.006	0.001	0.00000
N*toimiperhe41	-0.059	0.012	0.00004
N*graafinen ala	-0.076	0.017	0.00007
N*voimalaitokset	-0.054	0.019	0.00045
N*paperiteollisuus	-0.115	0.014	0.00000
N*kemianteollisuus	-0.069	0.013	0.00000
N*muut alat	-0.042	0.011	0.00023
N*lpoikkp	-0.088	0.025	0.00371
N*työtunnit/viikko	0.013	0.003	0.00002
<b>Selittäjien lukumäärä (myös vakio)</b>	<b>42</b>		
<b>Vapausasteet</b>	<b>6495</b>		
<b>Selitysaste</b>	<b>80.08%</b>		
<b>Korjattu selitysaste</b>	<b>79.95%</b>		
<b>Virheen hajonta</b>	<b>0.15085</b>		

**Huomautus:**

Kertoimien keskivirheet ja p-arvot on korjattu heteroskedastisuuden suhteen.

Mallissa PM on 42 selittäjää, jotka kuvaavat työtehtävän tyyppin, työtehtävän vaativuustason, iän, koulutuksen, toimialan, eräiden muiden tekijöiden (kunnan kalleusluokan ja työviikon pituuden) ja sukupuolen vaikutusta palkanmuodostukseen. Kaikki nämä selittäjät ovat luontevia ja niillä on tilastollisesti kiistaton merkitys palkanmuodostuksessa. Yhdessä ne selittävät suuren osan (80.08 %) henkilöiden välisestä palkkavaihtelusta. Malli PM kuvaa kaikkien toimihenkilöiden palkanmuodostusta. Jatkoanalyysin kannalta on havainnollista muodostaa erilliset estimoidut palkkayhtälöt naisten ja miesten osajoukossa. Tämä on tehty taulukossa 4.3.

**TAULUKKO 4.3.** *Miesten ja naisten palkkamallit (työnarvostusfunktiot)*

Muuttuja	Miesten mallin kerroin	Muutos	Naisten mallin kerroin
vakio	-0.482	-0.578	-1.060
kunnan kalleusluokka	-0.046		-0.046
ikä	-0.025		-0.025
kontvaat	0.057		0.057
teknvaat	0.106		0.106
koulutustaso	0.089		0.089
tutkinnon suoritusvuosi	-0.001		-0.001
työtunnit/viikko	-0.018	0.013	-0.005
lpoikkp	0.118	-0.088	0.030
yl31	-2.381	0.236	-2.145
ko21	-0.107		-0.107
ikä*yl31	-0.046	-0.006	-0.052
sqrt(koulutustaso)	-0.195		-0.195
sqrt(ikä)	0.356		0.356
sqrt(ikä)*yl31	0.698		0.698
akat. koulutus*yl31	0.100		0.100
lpoikkp*yl31	0.350		0.350
lpoikkp*ko21	0.215		0.215
toimiperhe12	-0.040		-0.040
toimiperhe14	-0.066		-0.066
toimiperhe21	0.027		0.027
toimiperhe41	0.049	-0.059	-0.009
graafinen ala	0.159	-0.076	0.082
voimalaitokset	0.108	-0.054	0.054
puhelinlaitokset	0.054		0.054
konsultit	0.039		0.039
sahateollisuus	-0.055		-0.055
paperiteollisuus	0.130	-0.115	0.016
rakennusteollisuus	0.074		0.074
kemianteollisuus	0.083	-0.069	0.014
muut alat	0.059	-0.042	0.017



Miesten ja naisten palkkamalleissa selittävät muuttujat ovat samoja, mutta osa niistä vaikuttaa eri tavalla palkanmuodostukseen sukupuolesta riippuen. Selittäjät, joiden vaikutus on erilainen miesten ja naisten kohdalla, ovat estimoidun palkkadiskriminaatiofunktion argumentteja. Niiden (estimoidut) kertoimet palkkadiskriminaatiofunktiossa ovat yhtä kuin (estimoidut) erot vaikutuksessa palkanmuodostukseen (taulukon 4.2 keskisarakkeen vastaluvut), jotka itse asiassa ovat naisindikaattorin ja kyseisten selittäjien interaktioiden kertoimien vastaluvut mallissa PM. Luvun 2 merkintöjä käyttäen logaritminen palkkadiskriminaatiofunktio on mallin PM tulosten perusteella

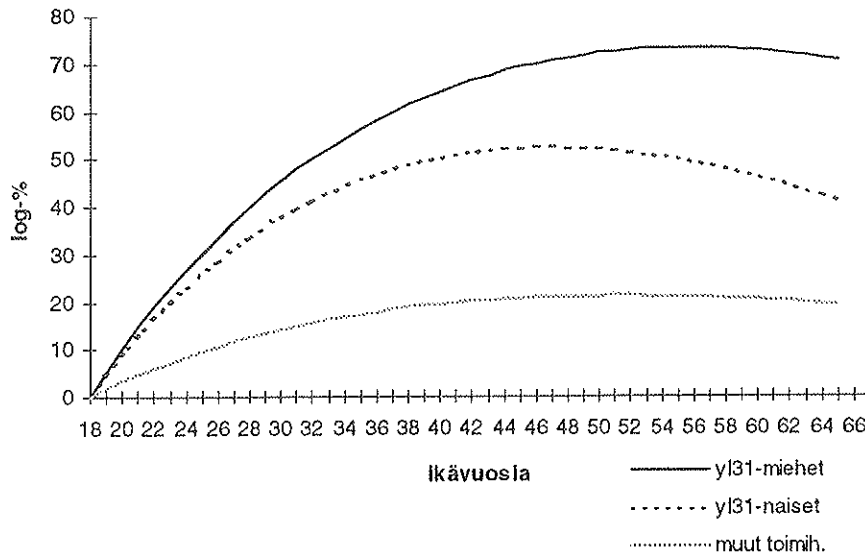
$$(4.1) \quad \text{LPDF}(\mathbf{X}) = 0.578 - 0.236 * y_{l31} + 0.006 * (y_{l31} * \text{ikä}) + 0.059 * \text{toimiperhe41} + \\ + 0.088 * \text{lpoikkp} - 0.013 * \text{työtunnit/viikko} + 0.076 * \text{graafinen ala} + \\ + 0.054 * \text{voimalaitokset} + 0.115 * \text{paperiteollisuus} + \\ + 0.069 * \text{kemianteollisuus} + 0.042 * \text{muut alat}.$$

(4.1) perusteella palkkadiskriminaatio riippuu toimihenkilöryhmästä, iästä, työtehtävästä ja toimialasta. Koska estimoidussa palkkamallissa selitettävä muuttuja on logaritminen poikkeama yleiseskipalkasta, palkkadiskriminaatiofunktio ilmoittaa estimoidun odotetun palkkadiskriminaation tiettyjen työn ominaispiirteiden kohdalla log-eroina ja sadalla kerrottuna log-prosenteissa. Seuraavaksi pyritään tulkitsemaan palkkadiskriminaatiofunktioita.

Palkkadiskriminaatiofunktion vakiotermin (mallin PM naisindikaattorin estimoidun kertoimen vastaluku) ja termi  $-0.236 * y_{l31}$  ovat tasosiirtymiä. Ne ovat teknisesti erittäin tärkeitä, mutta tässä kehikossa niillä ei ole itsenäistä tulkintaa palkkadiskriminaation kannalta. Ne eivät yksinään kerro mitään palkkadiskriminaatiosta ja sen vaihtelusta.

Estimoidun palkkadiskriminaatiofunktion perusteella ylempien toimihenkilöiden keskuudessa toimihenkilöiden ikä/palkka-profiili (palkkojen ikäprofiili) on erilainen naisten ja miesten kohdalla. Muissa toimihenkilöryhmissä palkkojen ikäprofiilit eivät tässä mallissa vaihtelee sukupuolen mukaan. Tämä ei tarkoita, että naisten ja miesten palkat kehittyisivät iän suh-

teen samalla tavalla, vaan sitä, että mallissa PM sellaiset erot on jo huomioitu ja selitetty muiden selittävien muuttujien avulla.



**KUVA 4.1.** Toimihenkilöiden palkkojen estimoidut ikäprofiilit

Kuvasta 4.1 ilmenee, että toimihenkilöiden palkat kasvavat aluksi nopeasti iän mukaan, erityisesti ylempien toimihenkilöiden osalta, mutta kääntyvät laskuun tietyn iän jälkeen. Konkaavit ikä/palkka-profiilit ovat inhimillisen pääoman teorian mukaisia. Käännekohta on ylempien miestoimihenkilöiden osalta noin 56 vuotta, kun ylempien naistoimihenkilöiden kohdalla palkat kääntyvät laskuun 47 ikävuoden jälkeen. Muilla toimihenkilöillä käännekohta on 51 ikävuoden kohdalla. Ylempien toimihenkilöiden keskuudessa samassa työtehtävässä työskentelevien naisten ja miesten välinen palkkaero on suurempi (miesten hyväksi) iäkkäämpien henkilöiden kohdalla. Suhteellinen palkkaero kasvaa, ceteris paribus, aina noin 0.6 log-prosenttia kun ikä lisääntyy yhdellä vuodella. Tätä kuvaa kuvion 4.1 kahden ylimmän käyrän kasvava erotus.

Palkkadiskriminaatiofunktiossa termi  $0.059 \cdot \text{toimiperhe41}$  tarkoittaa, että toimiperheessä 41, eli myyntityössä, naisten palkat ovat ylimääräisesti 5.9 log-prosenttia pienempiä kuin samaa työtä tekevien miesten palkat.

Termi  $0.088 \cdot \ln(\text{palkka})$  tarkoittaa, että palkkaerot kasvavat työtehtävän keskimääräisen palkkatason myötä. Kun siirrytään tietystä kolminumeroisesta työtehtävästä toiseen, jossa keskimääräinen palkkataso on 10 % korkeampi, on jälkimmäisessä työtehtävässä naisten ja miesten välinen palkkaero noin 0.85 log-% suurempi olettaen, että tarkastelun kohteena olevissa työtehtävissä toimihenkilöt ovat keskimäärin samanikäisiä, heillä on sama koulutustausta, ym. ominaisuudet.

Palkkadiskriminaatio on pienempi työtehtävissä, joissa säännöllinen viikkotyöaika on pitempi. Termi  $-0.013 \cdot \ln(\text{työtunnit/viikko})$  ilmoittaa, että työtehtävässä, jossa säännöllinen viikkotyöaika on tunnilla pitempi, sukupuolten välinen diskriminatiivinen palkkaero on keskimäärin 1.3 log-% pienempi, tietenkin ceteris paribus oletuksella.

Palkkadiskriminaatiofunktion termit  $0.076 \cdot \ln(\text{graafinen ala})$ ,  $0.054 \cdot \ln(\text{voimalaitokset})$ ,  $0.115 \cdot \ln(\text{paperiteollisuus})$ ,  $0.069 \cdot \ln(\text{kemianteollisuus})$  ja  $0.042 \cdot \ln(\text{muut alat})$  kuvaavat palkkadiskriminaation vaihtelua toimialoittain. Jos verrataan esimerkiksi keskenään kahta samaa työtä tekevää toimihenkilöryhmää (kussakin ryhmässä voi olla sekä miehiä että naisia), joista toinen on graafisella alalla ja toinen mallissa PM vertailukohtana olevalla toimialalla, metallialalla, palkkadiskriminaatio graafisen alan ryhmässä on keskimäärin 7.6 prosenttiyksikköä suurempi. Sen sijaan ei voida väittää, että yleensä palkkadiskriminaatio olisi graafisella alalla 7.6 prosenttiyksikköä korkeampi kuin metallialalla. Kuinka paljon suurempi (tai pienempi) palkkadiskriminaatio on graafisella alalla, verrattuna muiden toimialojen palkkadiskriminaatioon, riippuu myös siitä, miten eri toimialoilla työskentelevät naistoimihenkilöt jakautuvat toimihenkilöryhmien kesken ja työtehtäviin.

Yllä esitetyn palkkadiskriminaatiofunktion sanallisen kuvauksen tarkoituksena on antaa intuitiivinen käsitys siitä, mistä tekijöistä palkkadiskriminaation suuruus riippuu ja millä tavalla. Annetut tulkinnat perustuvat väistämättä ceteris paribus oletukseen, jolla ei aina ole konkreettista sisältöä.

Yhteenvetona todetaan, että palkkadiskriminaatio on suhteellisen monesta tekijästä riippuva suure. Estimoitu palkkadiskriminaatiofunktio on kuitenkin huomattavasti yksinkertaisempi lauseke kuin estimoitu palkkayhtälö, koska monen tekijän vaikutusta palkanmuodostukseen voidaan pitää samana miesten ja naisten kohdalla. Ei esimerkiksi löydy tilastollista näyttöä sille, että koulutustaso vaikuttaisi eri tavalla miesten ja naisten palkkoihin, joten koulutustason vaikutusta palkanmuodostukseen pidetään samana miesten ja naisten palkkayhtälöissä. Näin ollen koulutustaso ei ole myöskään estimoidun palkkadiskriminaatiofunktion argumentti. Seuraavassa luvussa arvioidaan palkkadiskriminaatiofunktion perusteella palkkadiskriminaation suuruutta toimihenkilöiden keskuudessa.

## 5 PALKKADISKRIMINAATION ARVIOINTI

### 5.1 NAISTEN JA MIESTEN VÄLISEN PALKKAERON DEKOMPOSITIO

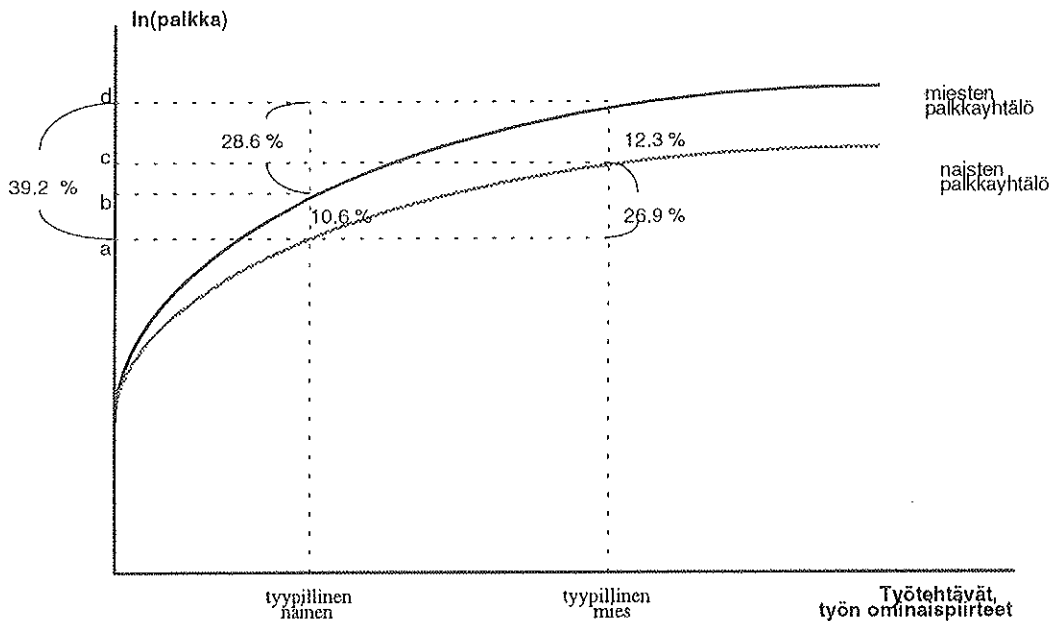
Luvussa 2 esitettiin Oaxacan (1973) kehittämä tapa jakaa sukupuolten välinen keskipalkkaero kahteen osaan, joista toinen tulkitaan miesten ja naisten työominaisuuksista ja työtehtäväeroista johtuvaksi ja toinen on varsinainen palkkadiskriminaatio. Huomattiin, ettei dekompositio ole yksikäsitteinen, vaan se voidaan suorittaa kahdella eri tavalla: Voidaan arvioida palkkadiskriminaatiota tyypillisen naisen kohdalla ja työtehtävistä ja henkilöominaisuuksista johtuva palkkaero-osa miesten palkkayhtälön mukaan, tai arvioida palkkadiskriminaatiota tyypillisen miehen kohdalla ja työtehtävistä johtuva palkkaero naisten palkkayhtälön mukaan

Tutkimusaineistossa naisten ja miesten välinen keskipalkkaero on 38.6 log-%, eli naisten palkka on keskimäärin noin 32 % pienempi kuin miesten palkka. Palkkamallin PM perusteella estimoitu palkkaero tyypillisen naisen ja tyypillisen miehen välillä on 39.2 log-%, eli hieman erilainen kuin keskipalkkaero, johtuen siitä, että palkkamallissa PM ikä ja koulutus-taso vaikuttavat epälineaarisesti palkanmuodostukseen (ks. edellä s.19). Taulukko 5.1 esittää tyypillisen miehen ja tyypillisen naisen välisen estimoidun palkkaeron Oaxacan dekomposition

**TAULUKKO 5.1.** *Tyypillisen naisen ja tyypillisen miehen välisen palkkaeron dekompositio*

<b>Estimoitu palkkaero</b>	39.2 log-%
<b>Miesten ja naisten työominaisuuksista ja työtehtävistä johtuva palkkaero</b>	
miesten palkkayhtälön mukaan	28.6 log-%
naisten palkkayhtälön mukaan	26.9 log-%
<b>Palkkadiskriminaatio</b>	
arvio tyypillisen naisen kohdalla	10.6 log-%
arvio tyypillisen miehen kohdalla	12.3 log-%

Tyypillisen naisen kohdalla palkkadiskriminaatio on 10.6 log-% ja tyypillisen miehen kohdalla arvio on 12.3 log-%, eli vajaat kaksi prosenttiyksikköä suurempi. Toisin sanoen, estimointitulosten perusteella vuonna 1993 teollisuustoimihenkilöiden keskuudessa naiset saavat keskimäärin 10.1 % pienempää palkkaa kuin samaa, naisille tyypillistä työtä tekevät miestoimihenkilöt. Palkkadiskriminaatio 12.3 log-% tyypillisen miehen kohdalla sen sijaan tarkoittaa, että miehet saavat keskimäärin 13 % isompaa palkkaa kuin samaa, miehille tyypillistä työtä tekevä naistoimihenkilö. Palkkadiskriminaation arvio tyypillisen naisen kohdalla antaa paremman käsityksen palkkadiskriminaatiosta, koska se kuvaa palkkadiskriminaatiota juuri niissä työtehtävissä, joita naiset todella tekevät. Kokonaispalkkaerosta palkkadiskriminaation osuus on 27 % eli reilu neljäsosa, jos sitä arvioidaan tyypillisen naisen kohdalla ja 31.4 %, jos sitä arvioidaan tyypillisen miehen kohdalla. Palkkadiskriminaatio toimihenkilöiden keskuudessa on näin ollen huomattavasti pienempi kuin miesten ja naisten välinen keskipalkkaero. Kuva 5.1 on tyypillisen naisen ja tyypillisen miehen välisen kokonaispalkkaeron dekomposition graafinen esitys (ks. myös edellä s. 20).



KUVA 5.1. Tyypillisen naisen ja tyypillisen miehen palkkaeron dekompositio

## 5.2 PALKKADISKRIMINAATIO TOIMIHENKILÖRYHMITTÄIN JA TOIMIALOITTAIN

Estimoitu palkkadiskriminaatiofunktio viittaa siihen, että palkkadiskriminaation suuruus vaihtelee melkoisesti toimihenkilöryhmittäin ja toimialoittain. On mielenkiintoista luoda laskennallisesti tietyille toimihenkilöryhmälle tai toimialalle tyypilliset nais- ja miespuoliset toimihenkilöt ja arvioida estimoidusta palkkadiskriminaatiofunktioista odotetun palkkadiskriminaation suuruus näiden kohdalla. Näin saadaan arvio keskimääräisestä palkkadiskriminaatiosta toimihenkilöryhmittäin ja toimialoittain. Taulukko 5.2 esittää estimaatit keskimääräisestä palkkadiskriminaatiosta toimihenkilöryhmittäin.

TAULUKKO 5.2: Keskimääräinen palkkadiskriminaatio toimihenkilöryhmittäin

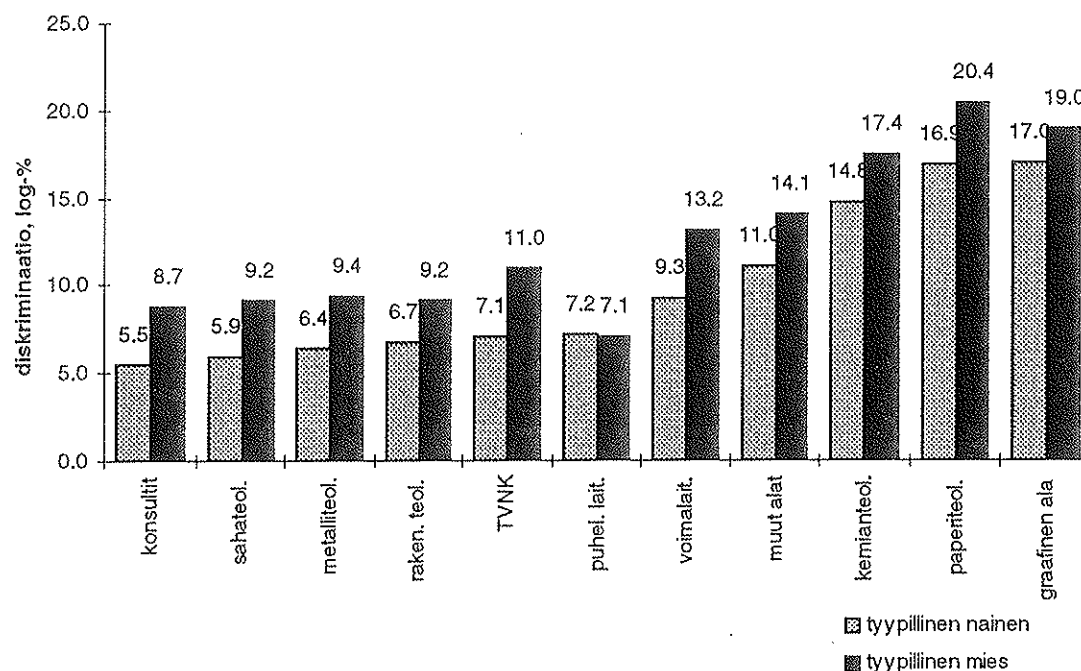
Arviointikohta	Diskriminaatio, log-%
<b>Konttoritoimihenkilöt</b>	
tyypillinen nainen	10.1
tyypillinen mies	13.4
<b>Tekniset toimihenkilöt</b>	
tyypillinen nainen	9.5
tyypillinen mies	9.7
<b>Ylemmät toimihenkilöt</b>	
tyypillinen nainen	14.1
tyypillinen mies	14.8

Palkkadiskriminaatio on suurin ylempien toimihenkilöiden keskuudessa, varsinkin naisten keskiarvon kohdalla mitattuna. Tyypillisen naispuolisen ylemmän toimihenkilön kohdalla palkkadiskriminaatio on 14.1 log-%, kun tyypillisen naispuolisen konttoritoimihenkilön kohdalla palkkadiskriminaatio on 10.1 log-% ja tyypillisen naispuolisen teknisen toimihenkilön kohdalla 9.5 log-%. Järjestyksen kannalta kokonaiskuva ei muutu, jos mitataan palkkadiskriminaatiota kullekin toimihenkilöryhmälle tyypillisen miehen kohdalla. Konttoritoimihenkilöiden osalta palkkadiskriminaatio tyypillisen miehen kohdalla on yli kolme prosenttiyksikköä suurempi kuin tyypillisen naisen kohdalla. Muissa ryhmissä taas tyypillisen

miehen kohdalla mitattu palkkadiskriminaatio on vain hieman suurempi kuin tyypillisen naisen kohdalla.

Koska ylemmät toimihenkilöt työskentelevät johtotehtävissä tai vaativissa asiantuntijatehtävissä (vaativa suunnittelutyö, lakiasiat yms.), voidaan saatujen tulosten perusteella sanoa, että naisjohtajat ja naisasiantuntijat kärsivät eniten palkkadiskriminaatiosta, vaikka he ovatkin paremmin palkattuja kuin naiset keskimäärin.

Seuraavaksi esitetään palkkadiskriminaatiota toimialoittain. Palkkadiskriminaatiofunktion perusteella on odotettavissa, että palkkadiskriminaatio on suurimmillaan paperialalla, graafisella alalla ja kemianteollisuudessa. Täytyy kuitenkin ottaa huomioon, että toimialoitteiset erot palkkadiskriminaatiossa riippuvat mm. myös siitä, miten naiset jakautuvat toimihenkilöryhmittäin toimialojen sisällä. Tarkempi käsitys toimialoitteisesta palkkadiskriminaatiosta syntyy arvioimalla sitä kunkin toimialan tyypillisessä nais- ja miestehtävässä. Palkkadiskriminaatio toimialoittain esitetään kuvassa 5.2.



KUVA 5.2. *Estimoitu palkkadiskriminaatio toimialoittain*



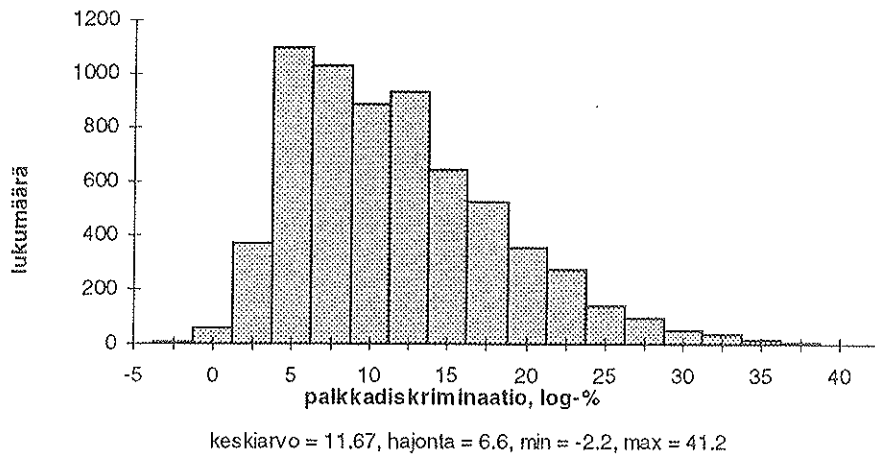
Palkkadiskriminaatio on suurimmillaan graafisella alalla ja paperiteollisuudessa. Kaikilla toimialoilla, puhelinlaitoksia lukuunottamatta, palkkadiskriminaatio tyypillisen miehen kohdalla on 2-3 prosenttiyksikköä korkeampi kuin tyypillisen naisen kohdalla Toimialan sisäinen palkkadiskriminaatio on siis yleisesti ottaen ainakin jossain määrin suurempi miesvaltaisissa tehtävissä. Jos järjestetään toimialat palkkadiskriminaation suuruuden mukaan, järjestys ei muutu olennaisesti riippumatta palkkadiskriminaation arviointikohdasta. Käytännössä estimoitua palkkadiskriminaatiota konsulttien ryhmässä, sahateollisuudessa, metalliteollisuudessa, rakennusteollisuudessa, TVNK:ssa ja puhelinlaitoksilla voidaan pitää lähes samana, ainakin naisille tyypillisissä tehtävissä. Pienet erot palkkadiskriminaatiossa näillä toimialoilla johtuvat pääasiassa siitä, että naiset jakautuvat kyseisillä aloilla hieman eritavalla toimihenkilöryhmien kesken. Tyypillisen naisen kohdalla mitattuna estimoitu palkkadiskriminaatio on suurimmillaan graafisella alalla, mikä on jossain määrin yllättävää, ottaen huomioon, että palkkadiskriminaatiofunktiossa paperiteollisuuden kerroin on 0.115, kun graafisen alan kerroin on vain 0.076. Tulos johtuu muutamasta tekijästä. Graafisella alalla peräti 33 % naistoimihenkilöistä työskentelee toimiperheessä 41 (myyntityö), jossa palkkadiskriminaatio on suuri, kun taas paperialalla vastaava luku on vain 9.5 %. Graafisen alan naistoimihenkilöiden säännöllinen viikkotyöaika on yli tunnin lyhyempi kuin paperialalla ja he myös sijoittuvat keskimäärin vähän paremmin palkattuihin tehtäviin. Estimoidun palkkadiskriminaatiofunktion mukaan nämä ovat palkkadiskriminaatiota kasvattavia seikkoja. Lopputulos on osoitus siitä, että toimialan kerroin palkkadiskriminaatiofunktiossa ei välttämättä yksin kerro, millä alalla palkkadiskriminaatio on suurempi ja kuinka paljon.

On myös tarkasteltu, ovatko palkkadiskriminaation toimialakohtaiset erot jollain tavalla sidoksissa toimialalla työskentelevien toimihenkilöiden sukupuolijakaumaan, eli riippuko toimialan palkkadiskriminaatio naistoimihenkilöiden osuudesta kaikista toimialalla työskentelevistä toimihenkilöistä. Sellaista riippuvuutta ei kuitenkaan havaittu. Esimerkiksi graafisella alalla naisten osuus toimihenkilöistä on suurin (62.3 %, sama kuin TVNK:ssa) ja myös palkkadiskriminaatio on suurimpia. Toisella varsin suuren palkkadiskriminaation alalla, eli paperiteollisuudessa, naistoimihenkilöiden osuus on 36.2 %, mikä on hyvin lähellä

keskimääräistä naisten osuutta toimihenkilöistä (36.9 %). Myös matalan diskriminaation aloilla naisten osuus toimihenkilöistä vaihtelee melkoisesti. Näin ollen on vaikeaa tarjota helppoa selitystä toimialoittaisen palkkadiskriminaation vaihtelusta. Saattaa olla, että se on jollain tavalla riippuvainen toimialan mies- tai naisvaltaisuudesta yleensä, eli kun naisvaltaisuus mitataan naisten osuudella kaikista alalla työskentelevistä (sekä työntekijät että toimihenkilöt). Asiaa ei ole tutkittu tässä yhteydessä sen enempää.

### **5.3 PALKKADISKRIMINAATION JAKAUMA TUTKIMUSAINEISTOSSA**

Kuten aikaisemmin todettiin, palkkadiskriminaatiofunktion avulla voidaan arvioida odotettu palkkadiskriminaatio kunkin toimihenkilön kohdalla, eli missä tahansa konkreettisesti työtilanteessa, jossa tietty toimihenkilö on ollut tilastointivaiheessa. Voidaan osoittaa (ks. tarkemmin Vartia ym. 1992, 82-83), että palkkadiskriminaatio tyypillisen naisen kohdalla on likimain sama kuin keskimääräinen palkkadiskriminaatio naisten osajoukossa. Jos palkkadiskriminaatiofunktio on lineaarinen argumenttiensa suhteen, yhtäsuuruus pätee täsmälleen. Vastaavasti palkkadiskriminaatio tyypillisen miehen kohdalla ja myös missä tahansa muualla, esimerkiksi palkkadiskriminaation keskiarvot eri toimialoilla, voidaan itse asiassa laskea suoraan henkilötasolta tilanteeseen sopivalla aggregoinnilla. Palkkadiskriminaation jakauma tutkimusaineistossa mahdollistaa myös muiden palkkadiskriminaatiota koskevien kysymysten käsittelyn. Olisi kiinnostavaa tietää muun muassa löytyykö toimihenkilöiden keskuudesta sellaisia ryhmiä, joissa naissukupuoli on palkan kannalta etu, eli löytyykö konkreettisia työtilanteita, joissa palkkadiskriminaatiofunktio saa selkeästi negatiivisia arvoja (negatiivinen palkkadiskriminaatio tarkoittaa tietysti, että naisilla on palkkaetua verrattuna samassa työtilanteessa oleviin miehiin). Palkkadiskriminaatiofunktion jakauma esitetään kuvassa 5.3.



KUVA 5.3. Palkkadiskriminaation jakauma vuoden 1993 poikkileikkausotoksessa

Estimoitu palkkadiskriminaatio vaihtelee otoksessa -2.2 log-prosentista 41.2 log-prosenttiin. Keskiarvo, 11.67 log-%, on luonnollisesti 10.6 log-% (palkkadiskriminaatio tyypillisen naisen kohdalla) ja 12.3 log-% (palkkadiskriminaatio tyypillisen miehen kohdalla) välillä, lähempänä 12.3 log-prosenttia, koska otoksessa on lähes kaksi kertaa enemmän miehiä kuin naisia. Estimoitu palkkadiskriminaatiofunktio ennustaa pienintä (jopa negatiivista) palkkadiskriminaatiota pääasiassa sellaisten varsin nuorten (alle 33-vuotiaat), suhteellisen matalapalkkaisten ylempien toimihenkilöiden joukossa, jotka eivät työskentele suuren diskriminaation toimialoilla eivätkä ole myyntitehtävissä. Tämän perusteella ei kuitenkaan voida väittää olevan työtilanteita, joissa naissukupuolesta olisi hyötyä palkan kannalta. Mieluummin voidaan todeta, että on joitakin harvinaisia työtilanteita, joissa palkkadiskriminaatio on nolla tai hyvin pieni.

Estimoidun palkkadiskriminaatiofunktion perusteella palkkadiskriminaatio on suurimmillaan iäkkäimpien toimihenkilöiden keskuudessa (48-50 vuotta ja yli), jotka tekevät myyntityötä ja työskentelevät toimialoilla, joilla palkkadiskriminaatio on suuri.

#### 5.4 YHTEENVETO

Edellä esitetyt tulokset viittaavat siihen, että suuri osa miesten ja naisten palkkaerosta johtuu työtehtävistä ja henkilökohtaisista ominaisuuksista. Tämä oli odotettavissa, koska kuten luvussa 3 todettiin, naistoimihenkilöt sijoittuvat selvästi matalapalkkaisiin tehtäviin, ovat keskimäärin nuorempia ja vähemmän koulutettuja. Tulokset ovat samansuuntaisia kuin Vartian ym. (1992) tutkimuksessa. Siinä on tutkittu palkkadiskriminaatiota v. 1990 metallisaha- ja paperiteollisuuden suuryrityksissä sekä työntekijöiden että toimihenkilöiden osalta. Vartian tutkimuksen mukaan palkkadiskriminaatio on toimihenkilöillä 8.9 log-% naistyötehtävissä ja miestyötehtävissä 14.3 log-%. Jos tässä estimoidun vuotta 1993 kuvaavan palkkamallin perusteella tutkitaan palkkadiskriminaatiota vain mainituilla aloilla, saadaan sen arvioksi tyypillisessä naistyötehtävässä 8.8 log-%, arvion ollessa miestyötehtävässä 11.1 log-%. Tyypillisen naistyötehtävän osalta tulokset ovat hämmästyttävän samanlaisia. Mieskeskiarvon kohdalla sen sijaan arvioissa on 3 prosenttiyksikön ero, mikä ei kuitenkaan ole kovin suuri. Molempien tutkimusten tulokset viittaavat siihen, että miesvaltaisissa työtehtävissä palkkadiskriminaatio on suurempi kuin naisvaltaisissa. Tässä mainittu palkkadiskriminaation piirre ei ole yhtä selkeä. Tutkimusten tulokset ovat samankaltaisia myös sen suhteen, että paperiteollisuudessa palkkadiskriminaatio on paljon suurempi kuin metalliteollisuudessa ja mekaanisessa metsäteollisuudessa (sahateollisuudessa).

Käsitys palkkadiskriminaatiosta ei muutu ollenkaan, jos analysoidaan palkkadiskriminaatiota palkkamallista, jossa selittäjänä on henkilön iän sijasta työkokemus. Malli PM1, jossa selittäjänä on työkokemus, ja siihen perustuvan palkkadiskriminaatioanalyysin keskeiset tulokset on esitetty liitteessä 3.

## 6 ERÄIDEN EI-HAVAITTAVIEN TEKIJÖIDEN MERKITYKSESTÄ PALKKADISKRIMINAATION ARVIOINNIN KANNALTA JA NIIDEN MALLINTAMINEN PANEELIAINEISTOSSA

### 6.1 EI-HAVAITTAVAT TEKIJÄT JA KYVYKKYYSHARHAONGELMA PALKKA- TUTKIMUKSESSA

Empiirisessä palkkatutkimuksessa palkkaeroja selitetään yleensä koulutuksella, iällä, työkokemuksella, työn vaativuudella ja joukolla muita relevantteja selittäjiä, kuten toimiala tai ammattiryhmään kuuluminen. Palkan suuruuteen vaikuttavat myös sellaiset tekijät kuten lahjakkuus, ahkeruus ja asennoituminen työhön, jotka ovat vaikeasti, jos ollenkaan mitattavissa. Näin ollen niistä ei tyypillisesti ole tietoa tutkimusaineistossa, joten tutkijan kannalta kyseiset tekijät ovat ei-havaittavia (*unobservable*). Analyysin kannalta ongelmana on, että lahjakkaammat ja ahkerammat ihmiset hankkivat itselleen todennäköisesti korkeamman koulutuksen ja sijoittuvat vaativampiin tehtäviin, joten lahjakkuuden, ahkeruuden ja muiden senkaltaisten henkilökohtaisten ominaisuuksien jättäminen pois palkkayhtälöstä johtaa siihen, että tavallinen pienimmän neliösumman (PNS) estimaattori yliarvioi koulutuksen ja työn vaativuutta kuvaavien muuttujien vaikutusta palkkaan. Englanninkielisessä kirjallisuudessa ongelma tunnetaan nimellä *ability bias*, eli *kyvykkyysarha*. Se on saanut osakseen suurta huomiota ilmeisesti ennen kaikkea siitä syystä, että koulutuksen tuotto (koulutuksen vaikutus palkkaan) on mielenkiinnon pääkohtia monissa inhimillistä pääomaa käsittelevissä empiirisissä tutkimuksissa. Todettakoon, ettei tämä sinänsä suoraviivainen ja luontevalta tuntuva ajatus positiivisesta kyvykkyysarhasta ilmeisesti ole kaikkien tutkijoiden hyväksymä. Griliches (1977) osoittaa, että tiettyjen ehtojen vallitessa koulutus ja virhetermi tavanomaisessa regressiomallissa saattavat olla negatiivisesti korreloituneita, jolloin PNS-estimaattori aliarvioi koulutuksen tuottoa.

Palkkadiskriminaatiotutkimuksessa ei-havaittavista muuttujista voi periaatteessa koitua kahdenlaista haittaa. Ensimmäinen johtuu palkkadiskriminaation määritelmästä ja se on

potentiaalisesti erittäin vakava ongelma. Diskriminatiiviseksi palkkaeroksi katsotaan se osa miesten ja naisten välille jäävästä kokonaispalkkaerosta, joka selitetään naisindikaattorilla ja sen interaktioilla, kun mallissa kontrolloidaan relevanttien työtehtävää ja työsuoritusta kuvaavien tekijöiden suhteen. Tutkimuksen laatu kärsii melkoisesti, jos mallista jää pois palkkaeroja selittävä muuttuja, jonka suuruus on keskimäärin erilainen miesten ja naisten kohdalla. Esimerkiksi jos miehet ovat keskimäärin naisia paremmin koulutettuja, ja koulutus ei ole mallin selittäjä, osa miesten ja naisten välisestä palkkaerosta, jota väitetään diskriminatiiviseksi, johtuu itse asiassa miesten korkeammasta koulutuksesta.

Ei ole perusteita ajatella, että samassa työtehtävässä, saman koulutuksen saaneet ja samankäiset (samaa työkokemusta omaavat) miehet ja naiset olisivat ahkeruudeltaan ja lahjakkuudeltaan keskimäärin erilaisia (väitehän olisi puhtaasti seksistinen). Sellaisten miesten ja naisten työsuorituksen voidaan päinvastoin odottaa olevan keskimäärin sama sukupuolesta riippumatta, joten lahjakkuuden ja ahkeruuden huomiotta jättäminen työnarvostusfunktiosta tuskin vääristää suoraan palkkadiskriminaation arviointia. Näin on tietenkin sillä ehdolla, että palkkamalli muuten kuvaa palkanmuodostuksen ja erityisesti siinä esiintyvät erot sukupuolten välillä oikein.

Kyvykkyysarha on toinen mahdollinen ongelma, joka saattaa esiintyä, jos lahjakkuuden ja ahkeruuden suhteen ei kontrolloida. Yleisesti ottaen kyvykkyysarhan ongelma ei rajoitu pelkästään koulutuksen tuoton estimaattiin, vaan myös kaikkien muiden koulutuksen kanssa korreloituneiden selittäjien kertoimet estimoidaan harhaisesti. Jos esimerkiksi koulutustaso on korreloitunut iän kanssa, mikä on luonteva ajatus - nuoremmat ihmiset ovat yleisesti ottaen paremmin koulutettuja - myös iän vaikutus palkkaan on jossain määrin väärin estimoitu poisjätettyjen selittäjien vuoksi. Palkkadiskriminaatiotutkimuksessa koulutuksen tuoton tarkka arviointi ei ole samalla tavalla kiinnostavaa kuin inhimilliseen pääomaan keskittyvissä tutkimuksissa, mutta mikäli koulutuksen tuotto on erilainen sukupuolten välillä, on tämän tuottoeron harhaton estimointi tärkeää. Sama pätee kaikkien muidenkin muuttujien osalta, jotka ovat palkkadiskriminaatiofunktion argumentteja. Jos voidaan

katsoa, että kyvykkyysarha on sama naisten ja miesten palkkayhtälöissä kaikkien palkkadiskriminaatiofunktion argumentteina olevien muuttujien osalta, kysymys on palkkadiskriminaation arvioinnin kannalta irrelevantti. Pelkästään PNS-estimaattorin ominaisuuksien perusteella päätellen näin ei tarvitse olla.

Onko lahjakkuuden ja ahkeruuden poisjättäminen palkkamalleista sitten vaarallista palkkadiskriminaatiotutkimuksessa? Empiirisessä palkkadiskriminaatiotutkimuksessa lahjakkuuden ja ahkeruuden vaikutusta palkanmuodostukseen ei yleisesti ottaen huomioida, koska käytetyt aineistot eivät anna siihen mahdollisuutta. Ajatus, että miehet ja naiset, jotka ovat samassa tehtävässä ja ovat yhtä koulutettuja ja kokeneita, pystyvät keskimäärin samanlaiseen työsuoritukseen, puoltaa voimakkaasti käsitystä, jonka mukaan lahjakkuuden ja ahkeruuden poisjättäminen palkkamallista ei voi johtaa suureen virheeseen palkkadiskriminaation arvioinnissa. Kyvykkyysarhan kautta tapahtuva palkkadiskriminaation väärä arviointi on mahdollinen, vaikkakin ongelma vaikuttaa hieman keinotekoiselta. Yllä esitetyt argumentit ovat kuitenkin paljolti heuristisia. Muuttuuko käsityksemme palkkadiskriminaatiosta, kun mallissa kontrolloidaan lahjakkuuden ja ahkeruuden suhteen, on viime kädessä empiirinen kysymys, ja sitä kannattaa tutkia mikäli mahdollista. Ei kuitenkaan ole perusteltua ajatella kyvykkyysarhan olevan palkkadiskriminaation kannalta niin keskeinen ongelma, ettei tutkimustulosten perusteella saisi tehdä palkkadiskriminaatiota koskevia johtopäätöksiä, jos sitä ei pystytä ratkaisemaan. Kun on käytettävissä paneeliaineisto, palkkamalleissa voidaan henkilökohtaisilla vakiotermeillä ottaa palkanmuodostuksessa huomioon eroja, jotka johtuvat ei-havaittavista tekijöistä, kuten lahjakkuudesta ja ahkeruudesta. Menetelmällä eliminoidaan myös mahdollinen kyvykkyysarha. Luvussa 6.2 esitetään idea tarkemmin ja tarkastellaan sen soveltuvuutta.

## 6.2 PANEELIAINEISTO JA VAKIOISTEN VAIKUTUSTEN MALLI

Ajatellaan, että tutkimusaineiston  $i$ :n henkilön (logaritmoitu) palkka  $y_i$  voidaan kuvata regressiomallilla:

$$(6.1) \quad y_i = \alpha_i + \mathbf{x}_i' \boldsymbol{\beta} + \varepsilon_i,$$

jossa  $\mathbf{x}_i$  on selittäjävektori ilman vakiotermiä ja  $\alpha_i$  on *henkilöspesifi vaikutus palkkaan* ja virhetermin oletetaan toteuttavan klassisen regressiomallin oletukset. Vektorin  $\mathbf{x}_i$  elementit ovat yleisten palkanmuodostusta ja palkkaeroja kuvaavien selittäjien (koulutus, työtehtävä ym.) arvot  $i$ :n henkilön kohdalla. Spesifikaatio poikkeaa klassisesta regressiomallista vain siinä, että yhteisen vakiotermin sijasta jokaisella henkilöllä on oma vakionsa (oma henkilöindikaattori). Henkilöspesifille vaikutukselle ei voida antaa kovin täsmällistä tulkintaa, mutta sen taustalla on joku kombinaatio kyseiselle henkilölle ominaisista ja *ajassa muuttumattomista* tekijöistä. On luontevaa ajatella, että nämä ovat henkilön synnynnäisiä ominaisuuksia, joita ei pystytä muuttamaan. Henkilöspesifistä vaikutuksesta palkkaan käytetään usein nimitystä *henkilökohtainen vaikutus (individual effect)*. Henkilöspesifissä vaikutuksessa palkkaan selvästi sisältyy muun muassa lahjakkuudesta ja ahkeruudesta johtuva osa. Jos jokaisesta henkilöstä on ainoastaan yksi havainto, mallia (6.1) ei voida estimoida, koska siinä on enemmän selittäjiä kuin vapausasteita. Oletetaan, että on käytettävissä paneelianeisto, eli jokaisesta otoksen henkilöstä on havaintoja vähintään kahdelta periodilta. Tällöin PNS-menetelmä soveltuu suoraan. Ekonometrisessa kirjallisuudessa malli (6.1) tunnetaan nimellä *vakioisten vaikutusten malli (fixed effects model)*, josta käytetään jatkossa lyhennettä VV-malli. Vakioisten vaikutusten mallin ja siihen liittyvän estimointimenetelmän esittää suhteellisen yksinkertaisesti esimerkiksi Greene (Greene, 1993, 465-469). Koska mallissa (6.1) on usein tuhansia henkilöindikaattoreita, estimointi suoritetaan tyypillisesti kahdessa vaiheessa. Olkoon paneelissa  $N$  henkilöä ja  $T$  periodia. Alaindeksillä  $i$  viitataan henkilöön ja alaindeksillä  $t$  viitataan havaintoperiodiin. Ensin estimoidaan  $\boldsymbol{\beta}$  tavallisella PNS:lla henkilöiden keskiarvopoikkeamista:



$$(6.2) \quad y_{it}^* = \beta' x_{it}^* + \varepsilon_{it}^*, \quad i = 1, \dots, N, \quad t = 1, \dots, T,$$

jossa henkilön  $i$  keskiarvopoikkeamat muuttujalle  $y$  ja  $x$ -vektorille ovat

$$(6.3) \quad y_{it}^* = y_{it} - \frac{1}{T} \sum_{t=1}^T y_{it}, \quad x_{it}^* = x_{it} - \frac{1}{T} \sum_{t=1}^T x_{it}.$$

Ensimmäisen estimointivaiheen tuloksena saadaan parametrivektorin  $\beta$  estimaatti  $\mathbf{b}$ . Keskiarvopoikkeamat eliminoivat henkilöspezifit vaikutukset. Toisessa vaiheessa henkilöspezifien vaikutusten estimaatit  $a_i$  saadaan henkilökeskiarvojen välisestä yhtälöstä

$$(6.4) \quad a_i = \bar{y}_i - \mathbf{b}' \bar{x}_i, \quad \bar{y}_i = \frac{1}{T} \sum_{t=1}^T y_{it}, \quad \bar{x}_i = \frac{1}{T} \sum_{t=1}^T x_{it}.$$

Tapauksessa, jossa havaintoperiodeja on vain kaksi ( $t=1,2$ ), parametrivektorin  $\beta$  estimointia voidaan yksinkertaistaa edelleen differensoimalla havainnot ja estimoimalla malli

$$(6.5) \quad y_{i2} - y_{i1} = \beta'(x_{i2} - x_{i1}) + (\varepsilon_{i2} - \varepsilon_{i1}), \quad i = 1, \dots, N$$

(ks. Chamberlain, 1978, 50). Esitystapa (6.5) tuo selvästi esiin menetelmän keskeisimpiä pulmia: VV-mallin luotettava estimointi edellyttää, että riittävän suuri osa kussakin selittäjässä on havaintoperiodien eikä havaintoyksiköiden välillä. Yhtälöstä (6.5) nähdään suoraan, että jos joku selittävä muuttuja saa saman arvon molemmilla havaintoperiodeilla kaikkien havaintoyksiköiden (henkilöiden) kohdalla, syntyy identifioituvuusongelma eikä kyseisen muuttujan kerrointa pystytä estimoimaan. Jos muuttujissa on vain vähäistä vaihtelua periodista toiseen, estimointi juuri ja juuri onnistuu, mutta malli on pahasti multikollineaarinen. Sen perusteella on mahdotonta suorittaa luotettavaa tilastollista päättelyä. Inhimillistä pääomaa käsittelevät tutkimukset, kuten Chamberlain (1978) ja Hausmann ym. (1981), joissa kyvykkyysarha pyritään eliminoimaan paneeliaineiston avulla, lähtevät yleisesti ottaen siitä, että esimerkiksi koulutus on aikuisten henkilöiden tapauksessa vakio ajassa. Niissä haetaan keinoja koulutuksen tuoton estimoimiseksi etsimällä sopivia instrumenttimuuttujia. Näissä tutkimuksissa VV-malli muodostaa vain yhden estimointivaiheen (ks. Hausmann ym., 1981). Ehdotetut menetelmät ovat sekä teoreettisesti että estimointiteknisesti varsin vaativia ja

luottavat erilaisiin oletuksiin muuttujien välisistä korrelaatioista. Kyseisissä tutkimuksissa on kiinnostuksen kohteena lähes yksinomaan koulutuksen tuoton harhaton estimointi eikä muita selittäjiä ole kovin paljon. Tässä tutkimuksessa VV-mallia sovelletaan suoraan toivoen, että aineiston sisältämä informaatio olisi riittävä. Jälkeenpäin ajatellen tämä vaikuttaa varsin rohkealta, mutta sitä ei voitu tietää etukäteen. Toisaalta muita keinoja ei oikeastaan ollut käytettävissäkään.

### **6.3 VAKIOISTEN VAIKUTUSTEN MALLIN SOVELTUVUUSONGELMAT**

Seuraavaksi tarkastellaan eräitä VV-mallin estimoinnin yhteydessä esiintyneitä ongelmia. Tärkeä kysymys on, voidaanko VV-mallin tulosten perusteella konstruoida palkkadiskriminaatiofunktio tai joku sen vastine, eli löytyykö myös VV-mallille tapa esittää samaa työtä tekevien miesten ja naisten palkkaerot. Koska käytetty paneeliaineisto on varsin lyhyeltä ajanjaksolta, törmätään vakavaan multikollineaarisuusongelmaan. Aineiston eikokeellinen luonne osoittautuu myös pulmalliseksi tulosten tulkinnan kannalta.

VV-mallissa ei saa olla muuttujia, jotka eivät vaihtelee saman henkilön kohdalla yli aikaperiodien, koska silloin henkilöindikaattorien vuoksi syntyy täysmultikollineaarisuusongelma (identifioituvuusongelma) eikä mallia voida estimoida. Esimerkkejä tällaisista muuttujista ovat sukupuoli- tai rotuindikaattorit. Tämä tarkoittaa, ettei naisindikaattoria voida sellaiseenaan käyttää selittäjänä. Tavanomaisen regressiomallin naisindikaattorin kertoimen vastaluku on palkkadiskriminaatiofunktion vakiotermi, eli lähtötasoero miesten ja naisten palkanmuodostuksessa (ks. kaava (2.13), s. 17). VV-mallissa naisindikaattorin sisältämä informaatio hajautuu kaikkien otoksen naisten henkilökohtaisiin vakiotermeihin eikä lähtötasoeroa miesten ja naisten palkanmuodostuksessa voida suoraan estimoida. Tätä eroa voitaneen kuitenkin kuvata seuraavalla tavalla.

Merkitään :

$$(6.6) \quad \begin{aligned} \bar{\alpha}_M & - \text{henkilökohtaisten vakiotermin keskiarvo miesten osajoukossa} \\ \bar{\alpha}_N & - \text{henkilökohtaisten vakiotermin keskiarvo naisten osajoukossa} \\ d & = \bar{\alpha}_N - \bar{\alpha}_M. \end{aligned}$$

VV-mallissa  $d$  olisi tietyissä olosuhteissa tavanomaisen regressiomallin naisindikaattorin kertoimen vastine ja sen vastaluku olisi siten palkkadiskriminaatiofunktion vakiotermi.

Yllä esitettyä voidaan perustella seuraavasti: VV-mallissa erot henkilökohtaisten indikaattoreiden kertoimien (henkilöspezifien vaikutusten) välillä voivat johtua toisaalta henkilöiden välisistä eroista ei-havaittavissa ominaisuuksissa ja toisaalta siitä, että näitä ominaisuuksia arvostetaan työmarkkinoilla eri tavalla eri henkilöiden kohdalla. Jos hyväksytään ajatus, että ei-havaittavat tekijät saavat keskimäärin samoja arvoja miesten ja naisten osajoukossa, niin ero henkilöspesifien vaikutusten keskiarvojen välillä johtuu pelkästään siitä, että kyseiset tekijät vaikuttavat eri tavalla miesten ja naisten palkkoihin. Se tulkitaan tässä kehikossa palkkadiskriminaatioksi. Kysymykseksi nousee, minkätyyppisten ei-havaittavien tekijöiden vaikutusta palkkaan heijastuu henkilökohtaisten vakioiden kertoimiin. Jos on onnistuttu mallintamaan hyvin yleisten tekijöiden vaikutusta palkkaan, voidaan ajatella, että erot henkilökohtaisten vakiotermin kertoimissa heijastavat lähinnä eroja lahjakkuudessa, ahkeruudessa ja perhetaustassa ja/tai eroja niiden vaikutuksessa palkkaan. Tällöin yllä esitetty argumentti on pätevä, koska voidaan katsoa, etteivät naiset ja miehet poikkea keskimäärin toisistaan ym. tekijöiden suhteen. Tässä liitetään henkilöindikaattoreiden kertoimien eroihin varsin voimakas tulkinta, joka ei ole oikea, jos ne heijastavat esimerkiksi myös ikä- tai työkokemuseroja. Jatkossa osoitetaan, että näin voi käydä, jolloin esitetyt perustelut eivät ole päteviä ja  $d$ :n tulkitseminen tavanomaisen regressiomallin naisindikaattorin kertoimen vastineeksi on kyseenalainen.

Tavanomaisessa regressiomallissa selittäjinä saattaa olla muun muassa periodi-indikaattoreita, jotka ilmaisevat mallin tasosiirtymää havaintoperiodista toiseen. Jos kuitenkin VV-malliin halutaan selittäjiksi myös T-1 kappaletta havaintoperiodi-indikaattoreita (ts. halutaan estimoida ns. kaksisuuntainen VV-malli), syntyy identifioituvuusongelma, jos mallin selittäjiin kuuluu esimerkiksi henkilön ikä. Iän muutos havaintoperiodista toiseen on väistämättä samanlainen kaikkien henkilöiden kohdalla. Tässä tapauksessa henkilöindikaattorit yhdessä havaintoperiodi-indikaattorien kanssa ovat täysin multikollineaarisia lineaarisen iän kanssa. Sama ongelma syntyy myös, jos selittäjänä on lineaarinen työkokemus, mikäli kaikki henkilöt ovat olleet koko tarkastellun aikavälin työssä.<sup>1</sup> Epäidentifioituvuus tarkoittaa, että mallissa on täysin turhia selittäjiä ja poistamalla niitä tai asettamalla niiden välille sopivia sideehtoja informaatiota ei hävitetä. Saattaa kuitenkin olla, ettei informaatiota voida välttämättä esittää palkkojen mallintamisen tarkoituksen kannalta mielekkäässä muodossa. Luonnolliselta tuntuu estimoida VV-malli ilman periodi-indikaattoria; kahden periodin tapauksessa olisi voinut olla korkeintaan yksi periodi-indikaattori. Kuitenkin täysin vastaavanlaiseen tulokseen päädytään helpommin, jos mallista jätetään pois lineaarinen ikä. Mutta kun näin menetellään, ei mallissa mukana olevan muuttujan, eli iän tai periodi-indikaattorin tulkinta eikä henkilökohtaisten indikaattorien tulkinta ole enää selvä. Ikäerot henkilöiden välillä ovatkin vakiot ajassa samoin kuin erot lahjakkuudessa tai ahkeruudessa. Nämä seikat viittaavat siihen, että VV-mallia on joissain tilanteissa vaikeaa, ehkä suorastaan mahdotonta, parametreja palkkadiskriminaatiotutkimuksen kannalta sopivalla tavalla. Aiheeseen palataan myös empiiristen tulosten esittelyn yhteydessä.

Multikollineaarisuusongelmaa tarkasteltiin jossain määrin jo edellä kohdassa 6.2. Teoreettisella tasolla ajatellaan, että palkanmuodostuksen kannalta keskeisiä tekijöitä ovat henkilön koulutus, työkokemus, työtehtävän vaativuus yms. Henkilöspesifit vaikutukset palkkaan

---

<sup>1</sup>Paneelissa olevien henkilöiden on täytynyt olla vuosina 1991-1993 työssä ainakin suurimman osan ajasta aineiston poimintatavan vuoksi. Tilastosta laskettu kokemusvuosien muutos ei toki ole kaikkien osalta sama, mutta tämä johtuu selvästi suurelta osin tilastointivirheistä (kokemuksen muutos negatiivinen tai suurempi kuin 3 vuotta). Viime kädessä on katsottu, että on parasta pitää kokemuksen muutos samana kaikkien toimihenkilöiden kohdalla.

oletetaan pieniksi, eli niiden lisääminen malliin on teoreettisessa mielessä "hienosäätöä". Kuitenkin VV-mallissa yleisten palkanmuodostukseen vaikuttavien tekijöiden regressioker-toimien estimointi perustuu vain kyseisten tekijöiden havaintoperiodien väliseen vaihteluun, eli tässä tapauksessa muutoksiin, jotka tapahtuvat kunkin henkilön koulutustasossa, työteh-tävässä ja sen vaativuustasossa jne. (ks. kaava (6.5), s.57). Muutoksia on tyypillisesti vain harvan henkilön kohdalla, etenkin kun havaintoperiodien lukumäärä on pieni ja yleisten tekijöiden regressiokertoimien  $\beta$  estimoinnissa VV-malli käyttää tehokkaasti vain näiden tekijöiden havaintoperiodien välistä vaihtelua (6.5). Koska ei-kokeellisessa paneeliaineistos-sa suurin osa selittäjien vaihtelusta on havaintoyksiköiden (henkilöiden) eikä havaintoperio-dien välillä, käy niin, että henkilöspezifit tekijät käytännössä selittävät suurimman osan palk-kaeroista. Tulos ei johdu varsinaisesti siitä, että henkilöindikaattorien määrä saattaa olla hy-vin suuri: Jos kaikkien muiden selittäjien vaihtelu olisi pääosin havaintoperiodien välillä tästä ei aiheutuisi ongelmaa. Siksi empiirisissä tutkimuksissa on usein suositeltavaa käyttää VV-mallin sijasta ns. satunnaisten vaikutusten (*random effects*) mallia (ks. tarkemmin Greene, 1993, 469-479). Tämä spesifikaatio kuitenkin olettaa, että henkilökohtaiset vaikutukset ovat korreloimattomia muiden mallin selittäjien kanssa (mt., 479), joten se ei ratkaise kyvykkyys-harhan ongelmaa. Satunnaisten vaikutusten spesifikaatio ei ilmeisesti ole käyttökelpoinen vaihtoehto, mikäli päämääränä on poistaa kyvykkyysarha selittäjien kertoimista.

Kuten empiiristen tulosten yhteydessä ilmenee, tässä käytetty tutkimusaineisto ei sisällä riit-tävästi informaatiota, jotta voitaisiin estimoida luotettavasti riittävän yleinen, toimihenkilöi-den palkanmuodostusta totuudenmukaisesti kuvaava VV-malli. Koulutustaso muuttuu vain vähän; kahden vuoden aikana koulutustasossa on tapahtunut muutoksia ainoastaan 38 hen-kilön kohdalla, jotka muodostavat vain 0.7 % paneelin henkilöistä.<sup>1</sup> Tämä tarkoittaa, ettei koulutuksen vaikutusta voi estimoida hyvin VV-mallista. Tilanne on vastaavantyyppinen monen muunkin sellaisen muuttujan osalta, joka on aikaisempien tulosten perusteella tärkeä

---

<sup>1</sup> Koska koulutustason muutoksia tutkittaessa törmättiin selkeisiin tilastointivirheisiin, on katsottu, että koulutustason muutos on todellinen vain siinä tapauksessa, että vuonna 1993 kyseisen henkilön kohdalla on merkitty tutkinnon suoritusvuodeksi vuosi 1991, 1992 tai 1993.

palkanmuodostuksen kannalta. Varsin harva aineiston toimihenkilö vaihtaa tarkasteluaikana esimerkiksi työtehtävää tai toimialaa.

Aineiston ei-kokeellinen luonne voi johtaa harhaisiin tuloksiin ja väärin tulkintoihin. Muutokset selittäjissä, joihin VV-mallien estimointi perustuu, eivät ole arvottuja. Toisin sanoen prosessi, joka generoi niitä, ei ole välttämättä rinnastettavissa koetilanteeseen, jossa esimerkiksi sattumanvaraisesti valituille henkilöille annetaan lisää koulutusta. Kun vain muutama paneelissa oleva suorittaa lisätutkinnon työn ohella, on loogista kysyä, voidaanko VV-mallin antama estimaatti lisätutkinnon vaikutuksesta palkkaan tulkita keskimääräiseksi lisätutkinnon vaikutukseksi palkanmuodostukseen. Lisätutkinnon suorittaneilla on saattanut olla poikkeuksellisen suuri insentiivi kouluttautua: He luultavasti tiesivät, että juuri heidän kohdallaan lisäkoulutus johtaisi varsin suureen palkkapreemioon. Jos näin on, VV-mallin tulokset ovat oikeita vain siinä mielessä, että ne kuvaavat oikein, mikä on ollut koulutuksen tuotto lisäkoulutusta saaneiden kohdalla. VV-mallista saatu estimaatti luultavasti yliarvioi koulutuksen keskimääräistä tuottoa, joka on yleensä kiinnostuksen kohteena, ja on tässä mielessä harhainen.

#### **6.4 EMPIIRISET TULOKSET**

Tutkimusaineistona on käytetty TT:n palkka-aineistosta poimittua paneelaineistoa vuosilta 1991 ja 1993, jota kuvattiin luvussa 3. Ensin estimointiin tavanomainen regressiomalli PNS-menetelmällä, palkkamalli OLS1, käyttäen samaa mallinvalintamenetelmää kuin aikaisemmin poikkileikkausaineiston analyysin yhteydessä.<sup>1</sup> Seuraava askel olisi estimoida VV-malli, jossa on (tietyillä ehdoilla, ks. edellä s. 58-60) samat selittäjät, ja vertailla tavanomaisen regressiomallin tuloksia VV-mallin tuloksiin. Osoittautui, että aineistossa vain erittäin harva toimihenkilö on vaihtanut tarkasteluaikana toimialaa, joten toimialaindikaattorien käyttö VV-mallissa ei ollut suotavaa. Tästä syystä toimialakohtaisten palkkaerojen mallintamista yksinkertaistettiin, mikä tapahtui siten, että toimialaindikaattorien sijasta selittäjänä käytet-

---

<sup>1</sup> Mallin estimointitulokset on esitetty liitteessä 4.

tiin muuttujaa **lulikps**, eli tietyn toimialan toimihenkilöiden keskipalkan logaritmin ja toimihenkilöiden yleiskeskipalkan logaritmin erotusta (ks. edellä s.27). Taulukossa 6.1 esitetään tavanomainen regressiomalli OLS2, jossa toimialaindikaattorit on korvattu muuttujalla **lulikps** ja vastaavasti VV-mallit VVM1 ja VVM2. Mallissa VVM1 OLS2:n selittäjäjoukosta on jätetty pois naisindikaattori N, ja vuoden 1993 indikaattori V93 identifioituvuusongelmien vuoksi. Mallissa VVM2 V93 on mukana selittäjänä, mutta vastaavasti ikä-muuttuja on jätetty pois. Nähdään, että mallien VVM1 ja VVM2 tulokset ovat yleisten palkanmuodostukseen vaikuttavien tekijöiden osalta täysin identtiset. Henkilöspesifien tekijöiden estimaatteja ei esitetä. Niitä ei välttämättä tarvitsekaan laskea (ks. kaavat (6.2) - (6.4), s. 57). Mallissa VVM2 vuosi-indikaattorin V93 kerroin on täsmälleen kaksi kertaa iän kerroin mallissa VVM1<sup>2</sup>, mikä on kaavan (6.5) perusteella aivan oikea tulos.

Kaikkien toimihenkilöiden kohdalla iän muutos vuosien 1991 ja 1993 välillä on 2 ja kaikkien toimihenkilöiden kohdalla V93:n muutos on 1. Iän kertoimen tulkinta mallissa VVM1 ja vastaavasti V93:n kertoimen tulkinta mallissa VVM2 on epäselvä. Nämä kertoimet heijastavat nyt toisaalta toimihenkilöiden iän lineaarista vaikutusta palkkaan ja toisaalta yleistä talouden kehitystä. Kun katsotaan kaavaa (6.4), eli henkilökohtaisten vakiotermin estimäättoria, huomataan, että myös estimoitujen henkilökohtaisten vakiotermin tulkinta hankaloituu: Malleissa VVM1 ja VVM2 henkilökohtaiset vakiotermit heijastavat nyt itse asiassa mm. toimihenkilöiden ikäeroja.

Kuten luvussa 6.3 todettiin, on vaikea sanoa, miten palkkadiskriminaatiota tulisi edes periaatteessa arvioida esitetyistä VV-malleista, eli mikä olisi tavanomaisesta regressiomallista estimoidun palkkadiskriminaatiodfunktion vastine. Esitettyjen VV-mallien parametroidit eivät ole mielekkäitä tutkimusongelman kannalta. Parempia parametroidivaihtoehtoja ei kuitenkaan ole tässä tilanteessa helppo löytää.

---

<sup>2</sup> Iän negatiivinen kerroin on odotettu ja looginen esitettyjen palkkamallien parametroidinnissa. Iän neliöjuuren kohdalla kertoimen looginen etumerkki on positiivinen. Näillä etumerkeillä palkan ikäprofiilit ovat odotettua muotoa.

Muuttuja	OLS2		VVM1		VVM2	
	Kerroin	p-arvo	Kerroin	p-arvo	Kerroin	p-arvo
vakio	-0.15246	-	-	-	-	-
ikä	-0.01846	0.00000	-0.05409	0.00000	-	-
sqrt(ikä)	0.27248	0.00000	0.56274	0.00000	0.56274	0.00000
ikä*yl31	-0.05878	0.00000	-0.05458	0.00000	-0.05458	0.00000
sqrt(ikä)*yl31	0.86706	0.00000	0.69100	0.00000	0.69100	0.00000
koulutustaso	0.08400	0.00000	0.25085	0.00977	0.25085	0.00977
sqrt(koulutustaso)	-0.18249	0.00007	-0.46382	0.23560	-0.46382	0.23560
akat. koulutus*yl31	0.10264	0.00000	-0.02075	0.39716	-0.02075	0.39716
tutkinnon suoritusvuosi yl31	-0.00132	0.00000	-0.00565	0.00189	-0.00565	0.00189
kont21	-2.91010	0.00000	-2.07140	0.00000	-2.07140	0.00000
lpoikkp	-0.07176	0.00000	-0.03828	0.01786	-0.03828	0.01786
lpoikkp*yl31	0.10872	0.00000	0.01179	0.63415	0.01179	0.63415
lpoikkp*kont21	0.33379	0.00000	0.04982	0.08833	0.04982	0.08833
lulikps	0.26805	0.00000	0.03508	0.31867	0.03508	0.31867
lulikps*kont21	0.61368	0.00000	0.13210	0.19333	0.13210	0.19333
kontvaat	-0.23407	0.00170	-0.07557	0.60758	-0.07557	0.60758
teknvaat	0.04895	0.00000	0.03928	0.00000	0.03928	0.00000
työtunnit/viikko	0.10911	0.00000	0.03507	0.00000	0.03507	0.00000
kunnan kalleusluokka	-0.01702	0.00000	-0.00412	0.00169	-0.00412	0.00169
toimiperhe12	-0.04797	0.00000	-0.00676	0.28589	-0.00676	0.28589
toimiperhe14	-0.05080	0.00000	-0.00032	0.96860	-0.00032	0.96860
toimiperhe21	-0.06799	0.00000	-0.01964	0.16000	-0.01964	0.16000
toimiperhe41	0.03752	0.00000	0.01571	0.04329	0.01571	0.04329
N, naisindikaattori	0.04502	0.00000	0.01534	0.14547	0.01534	0.14547
N*yl31*sqrt(ikä)	-0.47702	0.00000	-	-	-	-
N*yl31	-0.08446	0.00000	0.00839	0.71525	0.00839	0.71525
N*kont21	0.48000	0.00000	-0.12119	0.40976	-0.12119	0.40976
N*kontvaat	-0.06088	0.00000	-0.03174	0.23130	-0.03174	0.23130
N*teknvaat	0.01268	0.00724	-0.00721	0.29693	-0.00721	0.29693
N*lpoikkp	0.01958	0.00040	0.02835	0.03747	0.02835	0.03747
N*yl31*lpoikkp	-0.16015	0.00000	-0.03020	0.41288	-0.03020	0.41288
N*lulikps	0.17254	0.00034	-0.06912	0.18424	-0.06912	0.18424
N*toimiperhe41	-0.28074	0.00015	0.16199	0.31701	0.16199	0.31701
N*työtunnit/viikko	-0.04035	0.00047	-0.01705	0.27676	-0.01705	0.27676
N*työtunnit/viikko	0.01030	0.00000	0.00762	0.00038	0.00762	0.00038
V93, v.1993 ind.	-0.03868	0.00000	-	-	-0.10819	0.00000
muut selittäjät			5464 henkilöindi- kaattoria		5464 henkilöindi- kaattoria	
Muuttujien lukumäärä	36		5497		5497	
Vapausasteet	10892		5431		5431	
Selitysaste	79.40%		98.64%		98.64%	
Korjattu selitysaste	79.30%		97.26%		97.26%	
Virhehajonta	0.1518		0.0553		0.0553	

Mallin OLS2 p-arvot on korjattu heteroskedastisuuden

TAULUKKO 6.1. Mallien OLS2, VVM1 ja VVM2 estimointitulokset



Estimoidut VV-mallien tulokset ovat muutenkin, erittäin vakavan multikollineaarisuusiongelman takia, varsin epäluotettavia, jopa huonoja. Alle puolet esitetyistä kertoimien estimaateista on tilastollisesti merkitseviä edes viiden prosentin virhetasolla, muutaman kertoimen merkki vaihtuu ja usein kertoimien estimaatit mallissa OLS2 ja VV-malleissa ovat aivan eri suuruusluokkaa. Luonnollinen tapa lieventää multikollineaarisuusiongelmaa olisi tarkastella paneeliaineistoa pitkältä ajanjaksolta, jossa on useita havaintoperiodeja. Silloin tulee esiin muita hankaluuksia, esimerkiksi palkanmuodostus on luultavasti muuttunut kymmenen vuoden aikana. Asplundin (1993b) mukaan koulutuksen ja työkokemuksen merkitys palkanmuodostuksessa on 80-luvulla vähentynyt, joten palkkojen mallintamisessa tulisi ottaa huomioon myös erot tietyn tekijän vaikutuksessa eri ajankohtina.

VV-mallien tulosten tulkintaa vaikeuttaa edelleen se, että käytetty aineisto on ei-kokeellinen. Jos tarkastellaan yksityiskohtaisemmin koulutuksen estimoitua vaikutusta palkkaan, havaitaan, että mallin OLS2 mukaan lisäkoulutuksella on selkeä, mutta suhteellisen vaatimaton<sup>1</sup> positiivinen vaikutus palkkaan. Asteen korkeamman tutkinnon suorittaminen kasvattaa, ceteris paribus, palkkaa keskimäärin vain noin 4 % (poikkeuksena on akateemisen tutkinnon suorittaminen ylempien toimihenkilöiden keskuudessa, mikä tuo ylimääräisen kymmenen prosentin palkankorotuksen). VV-malleissa estimoidun lisätutkinnon tuotto on karkeasti jopa kolminkertainen ja tämä näyttää olevan täysin ristiriidassa *a priori* käsityksen kanssa positiivisesta kyvykkyysarhasta. Todellisuudessa tulos kuvaa asiallisesti ainoastaan lisäkoulutuksen vaikutusta niiden 38 henkilön palkkaan, jotka ovat vuosina 1991-1993 suorittaneet lisätutkinnon. Mitä luultavimmin lisätutkinnon suorittaneilla on kuitenkin ollut poikkeuksellisen suuri insentiivi kouluttautua, joten on väärin tulkita VV-mallien estimaatit keskimääräiseksi koulutuksen tuotoksi. Samasta syystä ei voi tehdä johtopäätöksiä sen suhteen, onko tavallisen PNS-menetelmän estimaateissa kyvykkyysarhaa ja mikä on sen suunta.

---

<sup>1</sup> Muistettakoon, että esitetyissä malleissa koulutuksen vaikutus kuvaa ainoastaan koulutuksen aikaansaamaa palkkapreemiota työtehtävän (ammatin) sisällä. Koulutus kasvattaa myös mahdollisuuksia sijoittua paremmin palkattuihin tehtäviin ja näin ollen sen kokonaisvaikutus palkkaan on paljon suurempi. Asian selvittäminen ei kuitenkaan ole tavoitteena tässä tutkimuksessa.

Erittäin vakava multikollinearisuongelma ja aineiston ei-kokeellinen luonne vievät tilastollisen analyysin epävarmalle pohjalle. Etu, joka voitaisiin saada mahdollisen kyvykkyysarhan eliminoinnista, on pieni siihen nähden, minkälaisia muita ongelmia VV-mallien käyttö toi esiin. Ainakaan palkkadiskriminaatiotutkimuksen kannalta VV-malli ei näytä olevan mielekäs tilastollinen spesifikaatio.

VV-mallissa on silmiinpistävää henkilöindikaattorien valtava määrä. On tuskin tarpeellista, että jokaisella henkilöllä olisi oma vakionsa mallissa. Hypoteesi, jonka mukaan joidenkin henkilöiden vakiotermin on sama, on periaatteessa testattavissa F-testillä, mutta käytännössä tehtävä on erittäin työläs ja tuskin suoritettavissa näin suuressa paneelissa. Olisi mielenkiintoista tutkia, miten pelkästään paneelin sisältämän informaation perusteella voitaisiin ryhmitellä toimihenkilöt henkilöspezifien vaikutusten suhteen homogeenisiin ryhmiin. Näin voitaisiin estimoida palkkamalleja, joissa henkilökohtaisten indikaattorien sijasta käytetään ryhmäindikaattoreita. Ryhmien lukumäärän tulisi olla paljon pienempi kuin henkilöindikaattorien lukumäärä. Tämän tapaisilla malleilla voisi tutkia, voidaanko säilyttää VV-mallin hyvät puolet ja samalla käyttää tehokkaammin selittäjien havaintoyksiköiden välistä vaihtelua niiden regressiokertoimien estimoinnissa. Aihe ei kuitenkaan kuulu tutkielman piiriin.

Lopuksi taulukossa 6.2 esitetään lyhyesti tavanomaisiin regressiomalleihin OLS1 ja OLS2 perustuvat arviot keskimääräisestä palkkadiskriminaatiosta vuonna 1993 ajanjaksolla 1991-1993 TT:n jäsenyritysten palveluksessa olleiden toimihenkilöiden keskuudessa. Tulokset ovat hyvin samansuuntaisia kuin luvussa 4 esitetyt arviot palkkadiskriminaatiosta, vaikka paneeliaineistosta estimoitu palkkadiskriminaatiofunktio on vähän erilainen kuin poikkeikkausaineistosta estimoitu. Varsinkin tyypillisen naisen kohdalla mallista OLS1 estimoitu palkkadiskriminaatio 10.8 log-% on miltei täsmälleen sama kuin edellisen luvun mallin PM antama arvio.

Arviointikohta	Palkkadiskriminaatio, log-%	
	OLS1:n mukaan	OLS2:n mukaan
<b>Konttoritoimihenkilöt</b>		
tyypillisen naisen kohdalla	10.45	9.75
tyypillisen miehen kohdalla	13.81	12.84
<b>Tekniset toimihenkilöt</b>		
tyypillisen naisen kohdalla	8.35	6.5
tyypillisen miehen kohdalla	7.18	6.2
<b>Ylemmät toimihenkilöt</b>		
tyypillisen naisen kohdalla	15.63	14.05
tyypillisen miehen kohdalla	14.87	14.31
<b>Keskimäärin</b>		
tyypillisen naisen kohdalla	10.82	9.8
tyypillisen miehen kohdalla	11.17	10.39

**TAULUKKO 6.2.** Palkkadiskriminaatio paneelin toimihenkilöillä vuonna 1993 mallien OLS1 ja OLS2 mukaan

Yleisestikin kokonaiskuva keskimääräisestä palkkadiskriminaatiosta ja palkkadiskriminaatiosta toimihenkilöryhmittäin on sama, kuin mikä saatiin analysoimalla vain vuoden 1993 poikkileikkausotosta. Eroja on lähinnä malliin OLS2 perustuvassa arviossa palkkadiskriminaatiosta teknisten toimihenkilöiden kohdalla. Mallin PM antamat arviot ovat 3 prosenttiyksikköä suurempia kuin OLS2:sta tyypillisen naisen kohdalla estimoidut ja 3.5 prosenttiyksikköä suurempia tyypillisen miehen kohdalla. Myös mallin OLS1 mukaan teknisten toimihenkilöiden kohdalla keskimääräinen palkkadiskriminaatio on suurempi kuin OLS2:n mukaan, mutta ero on pienempi (1-2 prosenttiyksikköä). Täytyy muistaa, että toisaalta tutkittava populaatio on nyt erilainen kuin luvussa 4 ja toisaalta se, että OLS2 on karkeampi malli kuin OLS1. Itse asiassa OLS2 antaa systemaattisesti vähän pienempiä arvioita keskimääräisestä palkkadiskriminaatiosta toimihenkilöryhmittäin kuin OLS1 ja pelkästään vuotta 1993 koskeva malli PM. Yleinen kuva palkkadiskriminaatiosta on joka tapauksessa sama kuin poikkileikkausaineiston analyysin pohjalta syntyvä käsitys.

## LOPUKSI

Tämän työn tulokset osoittavat, että vuonna 1993 teollisuustoimihenkilöiden keskuudessa on palkkaeroja samaa työtä tekevien miesten ja naisten välillä. Palkkadiskriminaatio ei ole kaikissa olosuhteissa sama, vaan sen suuruus vaihtelee mm. eri työtehtävissä, toimialoittain ja toimihenkilöryhmittäin. Tähän tulokseen päädyttiin estimoimalla palkkadiskriminaatiofunktio, joka on kompakti kuvaus siitä, mistä tekijöistä ja millä tavalla palkkadiskriminaatio on riippuvainen. Tyypillisissä naistehtävissä palkkadiskriminaatio on symmetrisesti ilmaistuna noin 10.6 log-%, eli tyypillinen naistoimihenkilö saa noin 10 % pienempää palkkaa kuin vastaavaa työtä tekevä miestoimihenkilö. Tyypillisessä miestehtävässä palkkadiskriminaatio on tutkimuksen mukaan jonkin verran suurempi, 12.3 log-%. Arviointitavasta riippuen palkkadiskriminaatio on 1/4 - 1/3 miesten ja naisten keskimääräisestä kokonaispalkkaerosta (noin 39 log-%). Loput kokonaispalkkaerosta selittyy erilaisilla työominaisuuksilla ja työtehtäväeroilla. Toimihenkilöryhmittäin tarkasteltuna palkkadiskriminaatio on suurimmillaan ylempien toimihenkilöiden keskuudessa, 14.1 - 14.8 log-prosenttia. Tämä viittaa siihen, että johto- ja asiantuntijatehtävissä olevat naiset kärsivät suhteellisesti eniten palkkasyrjinnästä. Graafisella alalla ja paperiteollisuudessa palkkadiskriminaatio on jopa noin 20 log-%. Näiltä osin tulokset ovat samansuuntaisia kuin Vartian ym.(1992) tutkimuksessa. Tässä tarkastelun kohteena on kuitenkin laajempi perusjoukko toimihenkilöitä, joten tuloksetkin voidaan tulkita yleisemmin.

Tulosten oikean tulkinnan kannalta on tärkeää muistaa ettei siinä oteta kantaa kysymykseen, miksi naiset ja miehet sijoittuvat työmarkkinoilla varsin erilaisiin tehtäviin siten, että naisille tyypilliset tehtävät ovat huonommin palkattuja, vaan työtehtäviin sijoittuminen käsitellään annettuna. On myönnettävä, että mitä luultavimmin nämä allokatiiviset erot johtuvat osittain naisiin kohdistuvasta syrjinnästä, mutta ilmiötä on mahdotonta selvittää tilastollisilla menetelmillä, eikä kysymys ole enää yksinomaan palkkadiskriminaatiosta, vaan yleisemmin naisten asemasta yhteiskunnassa.

Palkkadiskriminaatiota on tutkittu kansantaloustieteessä yleisesti hyväksytyllä, joskin varsin abstraktilla tilastollisella menetelmällä. Analyysin lähtökohtana oli olettaus, että jos palkkasyrjintää ei ole, miesten ja naisten palkkayhtälöt ovat kaikilta osin identtisiä. Oaxaca kiinnittää huomiota siihen, ettei tämä ole ongelmatonta (ks. Oaxaca 1973, 708). Tyypillinen empiirinen tulos on, että naisten ikä/palkka- tai kokemus/palkka-profiilit ovat matalampia kuin miesten. Tässä näin on selvästi asian laita ylempien toimihenkilöiden tapauksessa. Oaxaca (1993) huomauttaa, ettei tuloksen tulkinta ole yksiselitteinen. Seuraavat tulkintavaihtoehdot ovat mahdollisia: Miehet saavat korkeampaa tuottoa investoinnille työpaikkakoulutukseen (on-the-job training), naiset investoivat vähemmän työpaikkakoulutukseen tai molemmat väitteet ovat tosia. Jos naiset investoivat työpaikkakoulutukseen miehiä vähemmän, kuten voidaan väittää sillä perusteella, että naiset suunnittelevat perheystistä lyhyempää työuraa, ei ikä/palkka- tai kokemus/palkka-profiilien eroja tulisi tulkita kokonaan palkkadiskriminaatiota kasvattavaksi tekijäksi. Puolustukseksi Oaxaca toteaa, että vaikka naiset investoisivatkin vähemmän työpaikkakoulutukseen, tämä voi johtua myös siitä, että heillä on huonommat mahdollisuudet tehdä sellainen investointi. Yleisesti ottaen empiirisessä palkkadiskriminaatiotutkimuksessa, kuten myös tässä, erot estimoiduissa ikä/palkka- ja työkokemus/palkka-profiileissa on tulkittu kokonaan palkkadiskriminaatiosta johtuviksi.

Käytetyn tutkimusmenetelmän tulosten luotettavuus on yleisestikin riippuvainen siitä, kuinka runsaasti ja tarkkoja tietoja palkanmuodostukseen vaikuttavista tekijöistä käytetty tutkimusaineisto sisältää. Vaikka käytetty aineisto on parhaita olemassa olevia, ei se ole kaikilta osin täysin tyydyttävä. Luotettavaa tietoa esimerkiksi toimihenkilöiden työkokemuksesta ei ollut käytettävissä. Siitä huolimatta virhe palkkadiskriminaation arvioinnissa, joka on saattanut aiheutua siitä, ettei työkokemuksesta ollut riittävää tietoa, on mitä luultavimmin jäänyt hyvin pieneksi, koska työtehtäviin sijoittuminen, johon työkokemus voimakkaasti vaikuttaa, on tässä työssä otettu eksogeeniseksi ja työtehtäväerot on varsin yksityiskohtaisesti mallinnettu. Tutkimuksessa tarjottiin vaihteoisia mallin spesifikaatioita. Toisessa mallissa selittäjien joukossa oli ikä ja toisessa palkkaperusteena olevat työkokemus-

vuodet (ylempien toimihenkilöiden osalta potentiaaliset työkokemusvuodet). Mallista riippumatta tulokset olivat palkkadiskriminaation kannalta täysin samanlaisia.

Aina on mahdollista luetella palkanmuodostukseen potentiaalisesti vaikuttavia tekijöitä, joita ei ole tutkittu. Kysymys on, ovatko sellaiset puutteet kriittisiä tuloksen kannalta. Tässä on pohdittu, missä määrin lahjakkuuden ja ahkeruuden kaltaisten tekijöiden poisjättäminen palkanmuodostuksen tilastollisesta mallintamisesta on vaarallista palkkadiskriminaation arvioinnin kannalta. Vaikka *a priori* ei ollut vahvoja syitä ajatella, että lahjakkuuden ja ahkeruuden tapaisten tekijöiden poisjättäminen tilastollisista palkkamalleista heikentäisi ratkaisevasti palkkadiskriminaation analyysia, on tarkasteltu mahdollisuutta ottaa analyysissa huomioon tällaiset tekijät paneeliaineiston avulla. Käytettävissä oleva keino, vakioisten vaikutusten malli, osoittautui varsin ongelmalliseksi, joten tältä osin työ jäi tuloksetta.

## LÄHTEET

- ALLÉN, T. - LAAKSONEN, S. - KEINÄNEN, P. - ILMAKUNNAS, S. (1990): Palkkaa työstä ja sukupuolesta. Tutkimus palkkaeroista Suomessa. *Tilastokeskus, tutkimuksia 169*.
- ASPLUND, R. (1993a): Essays on human Capital and Earnings in Finland. ETLA, A18, Helsinki.
- (1993b): Teollisuuden toimihenkilöiden palkat ja inhimillinen pääoma. ETLA, B89, Helsinki.
- ASPLUND, R. - BARTH, E. - SMITH, N. - WADENSJÖ, E. (1993): The Male - Female Wage Gap in the Nordic Countries. Teoksessa WESTERGÅRD-NIELSEN, N. (toim): *Wage Differentials in the Nordic Countries*. North-Holland, Amsterdam.
- BRUNILA, A. (1990): Naisten ja miesten palkkaerot vuosina 1975 ja 1985. *Työväen taloudellinen tutkimuslaitos, tutkimuksia 30*.
- CHAMBERLAIN, G. (1978): Omitted Variable Bias in Panel Data: Estimating the Return to Schooling. *Annales de l'INSEE, 30/31, 49-82*.
- GREENE, W. (1993): *Econometric Analysis*. Second Edition. Macmillan Publishing Company. New York.
- GRILICHES, Z. (1977): Estimating the Return to Schooling: Some Econometric Problems. *Econometrica, 45, 1-22*.
- GUNDERSON, MORLEY (1989): Male - Female Wage Differentials and Policy Responses. *Journal of Economic Literature Vol. XXVII, 46-72*.

- HAUSMAN, J. - TAYLOR, W. (1981): Panel Data and Unobservable Individual Effects.  
*Econometrica*. Vol. 49, No.6, 1377-1398.
- HEMMILÄ, P (1989): Naisten ja miesten ansioerot. *TTT Katsaus* 17. vsk, 1/1989.
- KETTUNEN, J. - MARJANEN, R. (1992): Suomen Työnantajain Keskusliiton  
palkkatilastot. *ETLA:n keskustelualoitteita No 425*.
- MINCER, J. (1974): Schooling, Experience and Earnings. Columbia University Press for  
NBER, New York.
- OAXACA, R. (1973): Male - Female Wage Differentials in Urban Labour Markets.  
*International Economic Review*, Vol 14, No. 3, 693 - 709.
- PASANEN, JARKKO (1995): Syrjinnän talousteoria ja naisten palkkasyrjinnän suuruus  
Suomen metsä- ja metalliteollisuuden suuryrityksissä. Pro gradu-tutkielma. Helsingin  
Yliopiston valtiotieteellinen tiedekunta.
- STK (1991): Ilmoitukset toimihenkilötilastoa varten elokuulta 1991. Jäsentiedote 31/91.
- TÖRNQVIST L. - VARTIA, P. - VARTIA, Y. (1985): How Should Relative Changes be  
Measured? *The American Statistician*, 39, 43 - 46.
- VARTIA, Y. - KURJENOJA, J. (1992): Palkkadiskriminaatio. *Helsingin Yliopiston  
kansantaloustieteen laitoksen tutkimuksia, Nro 60:1992*.
- VARTIA, Y. (1993): Palkkadiskriminaatiofunktio. Työpaperi esitelmästä XV  
kansantaloustieteen päivillä Joensuussa.



- VARTIA, Y. (1995): Promootiokirjan esitelmä ja promoottorin puhe. *Helsingin Yliopiston kansantaloustieteen laitoksen esitelmiä ja kirjoituksia, 11/1995.*
- VARTIAINEN, J. (1992): Edustavan paneeliaineiston poimiminen STK:n palkkatilastosta: korjattu yksityiskohtainen suunnitelma, ETLA, työpaperi.
- WHITE, H. (1980): A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity. *Econometrica*, 48, 817 - 838.

LIITE 1

Palkkaskriminaatioanalyysissä käytetyt muuttujat ja niiden tunnusluvut v. 1993 poikkileikkauksineistossa

muuttuja	selvennys	keskiarvo	hajonta	minimi	maksimi
1	kunnan kalleusluokka				
2	N	1.4735	0.4993	1	2
3	N*kunnan kalleusluokka	0.3685	0.4824	0	1
4	työtunnit/viikko	0.5351	0.7629	0	2
5	N*työtunnit/viikko	38.3350	1.3753	32.5	40
6	teknvaat	13.9970	18.3390	0	40
7	kontvaat	0.1890	0.5964	-3	2
8	lpoikpp	0.0420	0.8763	-6	5
9	yl31	-0.0380	0.2682	-0.6905	0.8324
10	ko21	0.3372	0.4728	0	1
11	N*yl31	0.3098	0.4624	0	1
12	N*ko21	0.0531	0.2242	0	1
13	lpoikpp*ko21	0.2549	0.4358	0	1
14	lpoikpp*yl31	-0.0840	0.1700	-0.6905	0.5524
15	N*teknvaat	0.0643	0.1539	-0.4685	0.8324
16	N*kontvaat	0.0050	0.3253	-3	2
17	N*lpoikpp	-0.0228	0.7703	-6	5
18	N*lpoikpp*ko21	-0.0848	0.1811	-0.691	0.832
19	N*lpoikpp*yl31	-0.0793	0.1627	-0.691	0.552
20	koulutustaso	0.0054	0.0612	-0.4685	0.8324
21	koulutustaso*yl31	3.0517	1.1046	1	7
22	koulutustaso*ko21	1.3613	1.9849	0	7
23	sqrt(koulutustaso)	0.7887	1.2651	0	6
24	sqrt(koulutustaso)*yl31	1.7156	0.3290	1	2.646
25	sqrt(koulutustaso)*ko21	0.6723	0.9536	0	2.646
26	N*koulutustaso	0.4868	0.7428	0	2.449
27	N*koulutustaso*yl31	0.9885	1.4294	0	7
28	N*koulutustaso*ko21	0.2090	0.9155	0	7
29	N*sqrt(koulutustaso)	0.6423	1.1724	0	5
		0.5925	0.7984	0	2.646

LIITE 1, jatkoa

muuttuja	selvitys	keskiarvo	hajonta	minimi	maksimi
30	$N \cdot \sqrt{\text{koulutustaso}} \cdot y/31$	0.1043	0.4451	0	2.646
31	$N \cdot \sqrt{\text{koulutustaso}} \cdot \text{ko}21$	0.3986	0.6953	0	2.236
32	akat. koulutus	0.1184	0.3231	0	1
33	$N \cdot \text{akat. koulutus}$	0.0231	0.1502	0	1
34	akat. koulutus $y/31$	0.1112	0.3144	0	1
35	$N \cdot \text{akat. koulutus} \cdot y/31$	0.0185	0.1348	0	1
36	akat.koulutus*ko21	0.0063	0.0790	0	1
37	$N \cdot \text{akat. koulutus} \cdot \text{ko}21$	0.0044	0.0665	0	1
38	tutkinnon suoritusvuosi	77.8800	7.1664	45	93
39	$N \cdot \text{tutkinnon suoritusvuosi}$	28.5930	37.6610	0	93
40	ikä	40.6290	8.9483	18	67
41	ikä $y/31$	13.4560	19.4980	0	67
42	ikä*ko21	12.3010	19.0840	0	64
43	$\sqrt{\text{ikä}}$	6.3346	0.7083	4.243	8.185
44	$\sqrt{\text{ikä}} \cdot y/31$	2.1180	2.9952	0	8.185
45	$\sqrt{\text{ikä}} \cdot \text{ko}21$	1.9381	2.9232	0	8
46	$N \cdot \text{ikä}$	14.5560	19.8430	0	67
47	$N \cdot \text{ikä} \cdot y/31$	2.0667	8.9378	0	67
48	$N \cdot \text{ikä} \cdot \text{ko}21$	10.1050	17.9040	0	64
49	$N \cdot \sqrt{\text{ikä}}$	2.3003	3.0440	0	8.185
50	$N \cdot \sqrt{\text{ikä}} \cdot y/31$	0.3293	1.3995	0	8.185
51	$N \cdot \sqrt{\text{ikä}} \cdot \text{ko}21$	1.5934	2.7508	0	8
52	työkokemus	18.5260	9.4387	0	50
53	työkokemus $y/31$	6.3540	10.3460	0	50
54	työkokemus*ko21	5.5194	9.8974	0	44
55	$(\text{työkokemus})^2/100$	4.3227	3.8206	0	25
56	$[(\text{työkokemus})^2/100] \cdot y/31$	1.4740	3.0359	0	25
57	$[(\text{työkokemus})^2/100] \cdot \text{ko}21$	1.2841	2.8798	0	19.36
58	$N \cdot \text{työkokemus}$	6.5662	10.3740	0	48
59	$N \cdot \text{työkokemus} \cdot y/31$	0.9604	4.5602	0	48
60	$N \cdot \text{työkokemus} \cdot \text{ko}21$	4.6605	9.3476	0	44

## LIITE 1, jatkoa

muuttuja	selvennys	keskiarvo	hajonta	minimi	maksimi
61	$N \cdot [(\text{työkokemus}^2)/100]$	1.5071	3.0122	0	23.04
62	$N \cdot [(\text{työkokemus}^2)/100]^{1/31}$	0.2169	1.2634	0	23.04
63	$N \cdot [(\text{työkokemus}^2)/100]^{1/21}$	1.0909	2.6865	0	19.36
64	toimiperhe11	0.0060	0.0770	0	1
65	toimiperhe12	0.1758	0.3807	0	1
66	toimiperhe13	0.0343	0.1819	0	1
67	toimiperhe14	0.0413	0.1990	0	1
68	toimiperhe21	0.2059	0.4044	0	1
69	toimiperhe22	0.0774	0.2673	0	1
70	toimiperhe31	0.0259	0.1587	0	1
71	toimiperhe32	0.0203	0.1412	0	1
72	toimiperhe33	0.0113	0.1058	0	1
73	toimiperhe41	0.1216	0.3269	0	1
74	toimiperhe42	0.0061	0.0780	0	1
75	toimiperhe43	0.0070	0.0836	0	1
76	toimiperhe51	0.0101	0.1000	0	1
77	toimiperhe52	0.0263	0.1601	0	1
78	toimiperhe61	0.0040	0.0629	0	1
79	toimiperhe71	0.0179	0.1326	0	1
80	toimiperhe72	0.0200	0.1402	0	1
81	toimiperhe73	0.0298	0.1701	0	1
82	toimiperhe74	0.0199	0.1396	0	1
83	toimiperhe75	0.0609	0.2391	0	1
84	toimiperhe76	0.0245	0.1545	0	1
85	toimiperhe77	0.0132	0.1140	0	1
86	toimiperhe81	0.0037	0.0605	0	1
87	toimiperhe82	0.0043	0.0653	0	1
88	toimiperhe83	0.0023	0.0479	0	1
89	toimiperhe84	0.0190	0.1364	0	1
90	toimiperhe85	0.0076	0.0871	0	1
91	toimiperhe86	0.0037	0.0605	0	1
92	metalliteollisuus	0.3601	0.4801	0	1

LIITE 1, jatkoa

muuttuja	selvennys	keskiarvo	hajonta	minimi	maksimi
93	graafinen ala	0.0580	0.2337	0	1
94	voimalaitokset	0.0535	0.2251	0	1
95	puhelinlaitokset	0.0430	0.2028	0	1
96	konsultit	0.0343	0.1819	0	1
97	sahateollisuus	0.0281	0.1654	0	1
98	paperteollisuus	0.0939	0.2918	0	1
99	rakennusteollisuus	0.0401	0.1962	0	1
100	kemian teollisuus	0.1043	0.3057	0	1
101	TNVK	0.0248	0.1555	0	1
102	muut alat	0.1599	0.3665	0	1
103	N*toimiperhe11	0.0002	0.0124	0	1
104	N*toimiperhe12	0.0314	0.1743	0	1
105	N*toimiperhe13	0.0135	0.1153	0	1
106	N*toimiperhe14	0.0145	0.1197	0	1
107	N*toimiperhe21	0.0118	0.1079	0	1
108	N*toimiperhe22	0.0109	0.1037	0	1
109	N*toimiperhe31	0.0049	0.0698	0	1
110	N*toimiperhe32	0.0067	0.0818	0	1
111	N*toimiperhe33	0.0090	0.0946	0	1
112	N*toimiperhe41	0.0476	0.2129	0	1
113	N*toimiperhe42	0.0024	0.0494	0	1
114	N*toimiperhe43	0.0011	0.0327	0	1
115	N*toimiperhe51	0.0066	0.0808	0	1
116	N*toimiperhe52	0.0155	0.1234	0	1
117	N*toimiperhe61	0.0006	0.0247	0	1
118	N*toimiperhe71	0.0087	0.0930	0	1
119	N*toimiperhe72	0.0193	0.1375	0	1
120	N*toimiperhe73	0.0283	0.1658	0	1
121	N*toimiperhe74	0.0118	0.1079	0	1
122	N*toimiperhe75	0.0600	0.2374	0	1
123	N*toimiperhe76	0.0226	0.1488	0	1

LIITE 1, jatkoa

muuttuja	selvennys	keskiarvo	hajonta	minimi	maksimi
124	N*toimiperhe77	0.0132	0.1140	0	1
125	N*toimiperhe81	0.0012	0.0350	0	1
126	N*toimiperhe82	0.0009	0.0303	0	1
127	N*toimiperhe83	0.0009	0.0303	0	1
128	N*toimiperhe84	0.0179	0.1326	0	1
129	N*toimiperhe85	0.0043	0.0653	0	1
130	N*toimiperhe86	0.0029	0.0538	0	1
131	N*metalliteollisuus	0.0991	0.2989	0	1
132	N*graafinen ala	0.0361	0.1866	0	1
133	N*voimalaitokset	0.0167	0.1281	0	1
134	N*puhelinlaitokset	0.0203	0.1412	0	1
135	N*konsultit	0.0115	0.1065	0	1
136	N*sahateollisuus	0.0113	0.1058	0	1
137	N*paperiteollisuus	0.0340	0.1811	0	1
138	N*rakennusteollisuus	0.0112	0.1051	0	1
139	N*kemian teollisuus	0.0493	0.2164	0	1
140	N*TNVK	0.0156	0.1240	0	1
141	N*muut alat	0.0635	0.2439	0	1
142	lulikps	-0.0074	0.0509	-0.16631	0.078864
143	N*lulikps	-0.0049	0.0337	-0.16631	0.078864

## LIITE 2. Muuttujien ULI, kontvaat ja teknvaat muodostaminen

### 1. Muuttujan ULI muodostaminen TT:n alkuperäisistä sopimusaloista

ULI	selvennys	sopimusalojen koodit TT:n aineistossa
1	metalliteollisuus	40
2	graafinen ala	4
3	voimalaitokset	28
4	puhelinlaitokset	29
5	konsultit	31
6	sahateollisuus	51
7	paperiteollisuus	52
8	kemianteollisuus	69 (v.91), 63 (v.93)
9	rakennusteollisuus	60
10	TNVK (tekstiili-, nahka-, vaate- ja kenkäteoll.)	5,19,27,30
11	muut alat	kaikki muut

### 2. Muuttujan kontvaat muodostaminen

Yleissopimuksen vaativuuslohkot	Koodi TT:n aineistossa	Muuttujan kontvaat arvo
4C	043	5
4B	042	4
4A	041	3
3C	033	2
3B	032	1
3A	031	0
2C	023	-1
2B	022	-2
2A	021	-3
1C	013	-4
1B	012	-5
1A	011	-6

### 3. Muuttujan teknvaat muodostaminen

Teknisten palkkaryhmät	Koodi TT:n aineistossa	Muuttujan teknvaat arvo
A	6	2
B	5	1
C	4	0
D	3	-1
E	2	-2
F	1	-3

**LIITE 3. Malli PM1, jossa selittäjänä on työkokemus, ja siihen perustuva palkkadiskriminaatioanalyysi**

*I. MALLI PM1:n ESTIMOINTITULOKSET*

Muuttuja	Kerroin	Keskivirhe	p-arvo
vakio	0.846		
kunnan kalleusluokka	-0.045	0.004	0.00000
kontvaat, kont. tehtävien vaativuustaso	0.059	0.002	0.00000
teknvaat, tekn. tehtävien vaativuustaso	0.107	0.003	0.00000
koulutustaso	0.135	0.021	0.00000
tutkinnon suoritusvuosi	-0.002	0.000	0.00000
työtunnit/viikko	-0.018	0.002	0.00000
yl31, ylempien indikaattori	-0.146	0.020	0.00000
ko21, konttoritoimihenkilöiden indikaattori	-0.106	0.011	0.00000
sqrt(koulutustaso)	-0.339	0.062	0.00000
akat. koulutus*yl31	0.125	0.013	0.00000
työkokemus	0.007	0.001	0.00000
(työkokemus^2)/100	-0.011	0.002	0.00000
työkokemus*yl31	0.025	0.002	0.00000
[(työkokemus^2)/100]*yl31	-0.037	0.005	0.00000
lpoikkp	0.120	0.023	0.00000
lpoikkp*yl31	0.351	0.030	0.00000
lpoikkp*ko21	0.241	0.039	0.00000
toimiperhe12	-0.039	0.006	0.00000
toimiperhe14	-0.066	0.011	0.00000
toimiperhe21	0.029	0.007	0.00002
toimiperhe41	0.050	0.012	0.00001
graafinen ala	0.163	0.017	0.00000
voimalaitokset	0.109	0.010	0.00000
puhelinlaitokset	0.050	0.008	0.00000
konsultit	0.041	0.010	0.00006
			jatkuu



LIITE 3, jatkoa

<b>Muuttuja</b>	<b>Kerroin</b>	<b>Keskivirhe</b>	<b>p-arvo</b>
sahateollisuus	-0.054	0.010	0.00000
paperiteollisuus	0.128	0.010	0.00000
rakennusteollisuus	0.072	0.010	0.00000
kemianteollisuus	0.085	0.010	0.00000
muut alat	0.060	0.009	0.00000
N, naisindikaattori	-0.580	0.119	0.00000
N*y131	0.077	0.025	0.00213
N*toimiperhe41	-0.061	0.014	0.00002
N*graafinen ala	-0.082	0.019	0.00002
N*voimalaitokset	-0.057	0.015	0.00018
N*paperiteollisuus	-0.112	0.013	0.00000
N*kemianteollisuus	-0.073	0.013	0.00000
N*muut11	-0.036	0.011	0.00188
N*y131*työkokemus	-0.005	0.001	0.00006
N*lpoikkp	-0.097	0.031	0.00167
N*työtunnit/viikko	0.013	0.003	0.00002
<b>Selittäjien lukumäärä (myös vakio)</b>	42		
<b>Vapausasteet</b>	6495		
<b>Selitysaste</b>	79.9		
<b>Korjattu selitysaste</b>	79.8		
<b>Virheen hajonta</b>	0.1515		

**Huomautukset:**

- a. Ylempien toimihenkilöiden osalta työkokemusvuodet ovat potentiaalisia työkokemusvuosia.
- b. Kertoimien keskivirheet ja p-arvot on korjattu heteroskedastisuuden suhteen.

LIITE 3, jatkoa

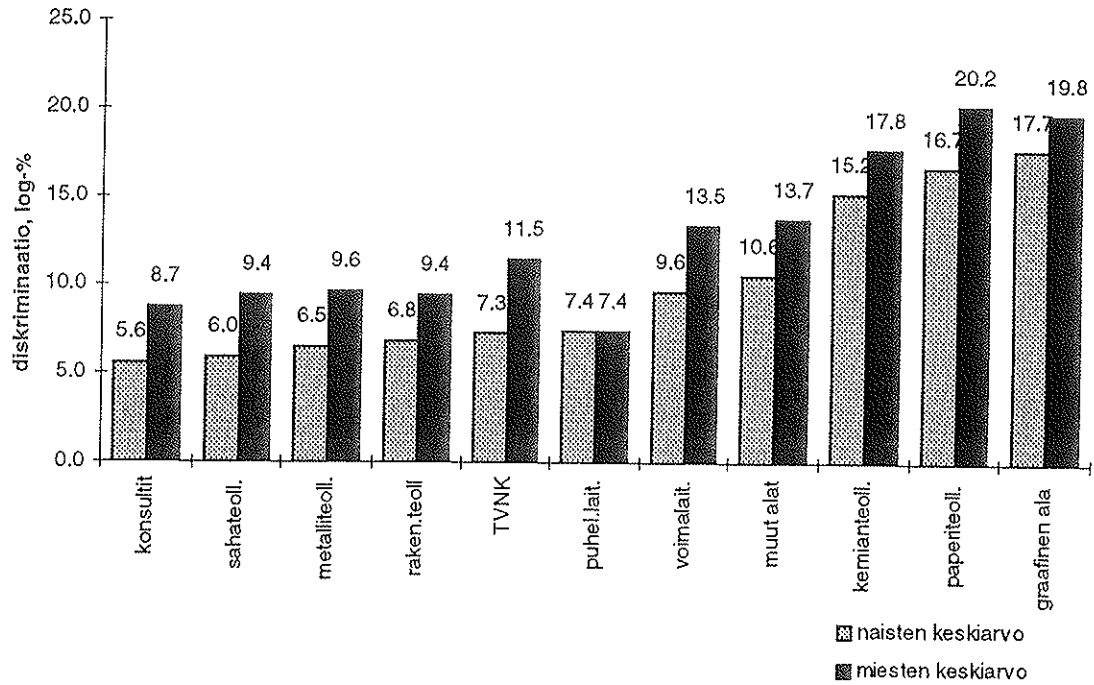
## **2. TYYPILLISEN NAISEN JA TYYPILLISEN MIEHEN VÄLISEN PALKKAERON DEKOMPOSITIO**

<b>Estimoitu palkkaero</b>	39.5 log-%
<b>Miesten ja naisten työominaisuuksista ja työtehtävistä johtuva palkkaero</b>	
miesten palkkayhtälön mukaan	28.8 log-%
naisten palkkayhtälön mukaan	27.1 log-%
<b>Palkkadiskriminaatio</b>	
tyypillisen naisen kohdalla	10.7 log-%
tyypillisen miehen kohdalla	12.4 log-%

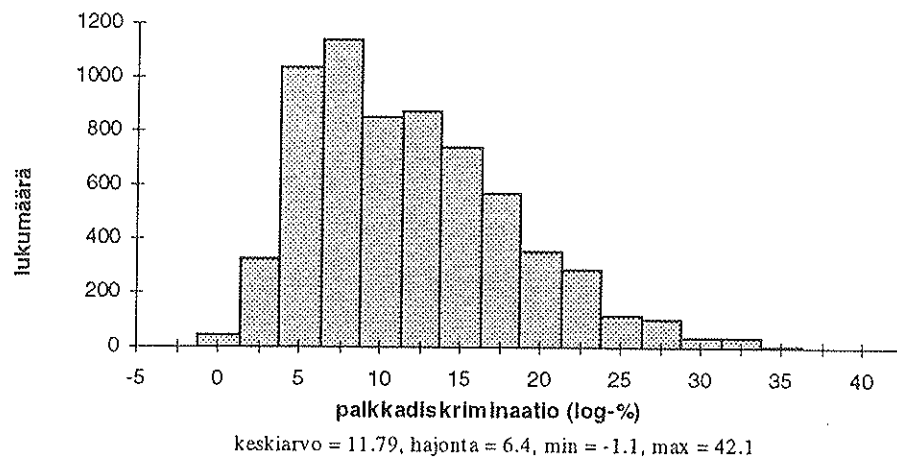
## **3. PALKKADISKRIMINAATIO TOIMIHENKILÖRYHMITÄIN**

<b>Arviointikohta</b>	<b>Diskriminaatio, log-%</b>
<b><i>Konttoritoimihenkilöt</i></b>	
tyypillinen nainen	10.2
tyypillinen mies	13.6
<b><i>Tekniset toimihenkilöt</i></b>	
tyypillinen nainen	9.6
tyypillinen mies	9.9
<b><i>Ylemmät toimihenkilöt</i></b>	
tyypillinen nainen	14.4
tyypillinen mies	14.8

**3. PALKKADISKRIMINAATIO TOIMIALOITTAIN**



**4. PALKKADISKRIMINAATION JAKAUMA V. 93 POIKKILEIKKAUSAINEISTOSSA**



ELINKEINOELÄMÄN TUTKIMUSLAITOS (ETLA)  
THE RESEARCH INSTITUTE OF THE FINNISH ECONOMY  
LÖNNROTINKATU 4 B, FIN-00120 HELSINKI

---

Puh./Tel. (09) 609 900  
Int. 358-9-609 900

Telefax (09) 601753  
Int. 358-9-601 753

KESKUSTELUAIHEITA - DISCUSSION PAPERS ISSN 0781-6847

- No 540 JULIANNA BORSOS - MIKA ERKKILÄ, Foreign Direct Investment and Trade Flows Between the Nordic Countries and The Baltic States. 29.09.1995. 43 p.
- No 541 RITA ASPLUND, The Gender Wage Gap in Finnish Industry in 1980-1994. An empirical analysis of non-manual workers. 11.10.1995. 28 p.
- No 542 TOR ERIKSSON - SUSANNA FELLMAN, Determinants of Firms' Operating Times - Some Evidence from Firm-Level Data. 23.10.1995. 19 p.
- No 543 TARMO VALKONEN, Corporate and Capital Income Tax Reform in a Numerical Overlapping Generations Model: The Case of Finland. 12.12.1995. 28 p.
- No 544 REIJA LILJA, Career Mobility in Finnish Industry. 14.12.1995. 22 p.
- No 545 JUKKA LASSILA - TARMO VALKONEN, Policy Credibility in Numerical Overlapping Generations Models. 28.12.1995. 28 p.
- No 546 EIJA KAUPPI - JUKKA LASSILA - TIMO TERÄSVIRTA, Short-Term Forecasting of Industrial Production with Business Survey Data: Experience from Finland's Great Depression. 26.01.1996. 20 p.
- No 547 JIAN-GUANG SHEN, FDI, Knowledge Spillover and Economic Growth in East Asia. 31.01.1996. 70 p.
- No 548 MAARIT SÄYNEVIRTA - PEKKA YLÄ-ANTTILA, Integraation ja yritysten kansainvälistyminen - vaikutuksia kotimaan talouteen. 01.02.1996. 58 s.
- No 549 LAURA PAIJA - PEKKA YLÄ-ANTTILA, The Impact of Structure and Competition on Employment in the Telecommunications Cluster Case Finland. 12.02.1996. 34 p.
- No 550 PENTTI VARTIA - PEKKA YLÄ-ANTTILA, Technology Policy and Industrial Clusters in a Small Open Economy - The Case of Finland. 16.02.1996. 15 p.
- No 551 PONTUS BRAUNERHJELM - PER HEUM - PEKKA YLÄ-ANTTILA, Internationalization of Industrial Firms. Implications for growth and industrial structure in the Nordic countries. 16.02.1996. 33 p.
- No 552 REIJO MANKINEN, Alkoholiveron alentamisen kansantaloudellisia vaikutuksia. 26.02.1996. 42 s.

- No 553 RITA ASPLUND, Koulutus, työura ja palkkaerot. 22.03.1996. 13 s.
- No 554 MARIANNE PAASI, The Absorptive Capacities of Estonian Firms. - Can a Technology-based Industrial Strategy Succeed? 22.03.1996. 17 p.
- No 555 HANNU HERNESNIEMI, Barriers to Economic Cooperation of Baltic Rim Countries. 10.04.1996. 46 p.
- No 556 ANNICK LARUELLE - MIKA WIDGRÉN, Is the Allocation of Voting Power among the EU States Fair? 17.04.1996. 19 p.
- No 557 JARI HYVÄRINEN, A Survey of Corporate Governance - Which Model for Transition Countries? 13.05.1996. 32 p.
- No 558 PASI KUOPPAMÄKI, Joint Implementation ilmastopolitiikan välineenä: Suomi ja lähialueiden kasvihuonekaasujen päästöjen rajoittaminen. 12.06.1996. 35 s.
- No 559 MIKA PAJARINEN, Työnantajan kansaneläkemaksun porrastuksen toimivuus vuosina 1993 ja 1994. 20.6.1996. 34 s.
- No 560 OKKO-PEKKA SALMIMIES, EU:n pakolais- ja maahanmuuttoasioita koskevan yhteistyön kehittäminen ja taloudellinen ulottuvuus. 10.07.1996. 32 s.
- No 561 AIJA LEIPONEN, Education, Tenure and Innovation in Manufacturing firms. 16.08.1996. 26 p.
- No 562 AIJA LEIPONEN, Education and Innovative Capabilities. 16.08.1996. 20 p.
- No 563 AIJA LEIPONEN, Competences, Innovation and Profitability of Firms. 16.08.1996. 16 p.
- No 564 JYRKI RUUTU, Suomen valuuttakurssijärjestelmä osana eurooppalaista järjestelmää. Historiallinen katsaus keskiajalta autonomian ajan loppuun. 29.08.1996. 22 s.
- No 565 HEIKKI PALM, Eläkeuudistuksen vaikutukset Suomen kansantalouden numeerisessa limitäisten sukupolvien mallissa. 02.09.1996. 34 s.
- No 566 JYRKI ALI-YRKKÖ, Teknologiaintensiivisten yritysten kansainvälistyminen - vaikutuksia Suomen talouteen. 05.09.1996. 53 s.
- No 567 VENLA SIPILÄ, Suomen ja Venäjän välinen kauppa 1986-1995. 06.09.1996. 100 s.
- No 568 EUGEN KOEV, Palkkadiskriminaatio teollisuuden toimihenkilöillä. 11.09.1996. 73 s.

Elinkeinoelämän Tutkimuslaitoksen julkaisemat "Keskusteluaiheet" ovat raportteja alustavista tutkimustuloksista ja väliraportteja tekeillä olevista tutkimuksista. Tässä sarjassa julkaistuja monisteita on mahdollista ostaa Taloustieto Oy:stä kopiointi- ja toimituskuluja vastaan hintaan.

Papers in this series are reports on preliminary research results and on studies in progress. They are sold by Taloustieto Oy for a nominal fee covering copying and postage costs.

d:\ratapalo\DP-julk.sam/11.09.1996