

VATT-KESKUSTELUALOITTEITA
VATT DISCUSSION PAPERS

309

PALKKADISKRIMI-
NAATIO SUOMEN
TEOLLISUUS-
SEKTORIN TOIMI-
HENKILÖILLÄ
VUONNA 2000

Antti Luukkonen

ISBN 951-561-456-2

ISSN 0788-5016

Valtion taloudellinen tutkimuskeskus

Government Institute for Economic Research

Hämeentie 3, 00530 Helsinki, Finland

Email: antti.luukkonen@vatt.fi

Oy Nord Print Ab

Helsinki, kesäkuu 2003

Saatteeksi

Tämä tutkimus on tarkastettu ja hyväksytty Helsingin yliopistossa pro gradu -tutkielmana huhtikuussa 2003. Kiitän ohjaajaani Yrjö Vartiaa monista pitkistä ja hedelmällisistä keskusteluista, joita kävimme asian tiimoilta.

Kiitän lämpimästi Teollisuuden ja Työnantajain Keskusliittoa aineiston käyttöluvasta ja VATT:ta mahdollisuudesta käyttää VATT:n resursseja tutkielman teon aikana talvella 2002–2003. Erityiskiitokseni esitän ekonomisti Tomi Kyyrälle ja tutkija Ossi Korkeamäelle, joilta sain useita hyödyllisiä kommentteja ja arvokasta palautetta työn edetessä.

Kaikista tutkimuksessa mahdollisesti esiintyvistä puutteista vastaan luonnollisesti itse.

Helsingissä 3.6.2003

Antti Luukkonen

LUUKKONEN ANTTI: PALKKADISKRIMINAATIO SUOMEN TEOLLISUUSSEKTORIN TOIMIHENKILÖILLÄ VUONNA 2000. Helsinki, VATT, Valtion taloudellinen tutkimuskeskus, Government Institute for Economic Research, 2003, (C, ISSN 0788-5016, No 309). ISBN 951-561-456-2.

Tiivistelmä: Tutkimuksessa selvitetään, esiintyykö Suomen teollisuussektorilla mies- ja naistoimihenkilöiden välillä palkkaeroja samassa työssä eli palkkadiiskriminaatiota. Sama työ määritellään hyvin tiukasti eli lähes 300 muuttujan avulla, ja samalla johdetaan uusi tapa ottaa huomioon toimihenkilöiden palkanmuodostuksen kannalta keskeiset työn vaativuusluokitukset.

Suuruudeltaan palkkadiiskriminaatio on tyypillisessä toimihenkilötehtävässä noin 8 log-prosenttia, joka on aiempia palkkaerotutkimuksia selvästi alempi luku. Palkkadiiskriminaatio riippuu yli 20 tekijästä, joista sukupuolten erilainen ikäkohtelu on selvästi voimakkain yksittäinen syrjintää aiheuttava palkanmuodostuksen piirre. Miehet ja naiset aloittavat työuransa likimain samasta tilanteesta, mutta iän tuoma työkokemus kasvattaa miesten palkkaa huomattavasti nopeammin kuin naisten.

Asiasanat: Palkanmuodostus, Palkkaerot, Syrjintä, Työn vaativuusluokitus, Teollisuus

LUUKKONEN ANTTI: PALKKADISKRIMINAATIO SUOMEN TEOLLISUUSSEKTORIN TOIMIHENKILÖILLÄ VUONNA 2000. Helsinki, VATT, Valtion taloudellinen tutkimuskeskus, Government Institute for Economic Research, 2003, (C, ISSN 0788-5016, No 309). ISBN-951-561-456-2.

Abstract: This study investigates whether white-collar men and women working in similar job are paid unequally in the industrial sector. ‘Similar job’ is defined very narrowly using almost 300 variables. A new way to measure the job complexity level – a crucial determinant of the wage of a white-collar worker – is introduced.

Wage discrimination in a typical job of the white-collar workers is some 8 logarithmic percent. Wage discrimination depends on more than 20 factors. The most important factor is the fact that work experience gained through aging contributes stronger to the wages of men than to the wages of women. In the beginning of the career men and women are paid equally.

Key words: Wage formation, Wage differences, Discrimination, Job complexity level, Industry

Yhteenveto

Miesten ja naisten välinen noin parinkymmenen prosentin suuruinen keskipalkkaero on osoittautunut teollisuusmaissa melko pysyväksi ilmiöksi, ja sitä on tutkittu paljon. Vaikka naisten osallistumisella työelämään on Suomessa pitkät perinteet ja vaikka yhteiskuntamme on monin tavoin tasa-arvoistunut ajan saatossa, ei palkkaero ole kadonnut maastamme mihinkään. Tämä on herättänyt vilkasta yhteiskunnallista keskustelua ilmiön syistä.

Tässä tutkimuksessa keskitytään palkkadiskriminaatioon, joka on yksi sukupuolten väliseen palkkaeroon vaikuttavista prosesseista; työssä selvitetään, esiintyykö Suomen teollisuussektorilla mies- ja naistoimihenkilöiden välillä palkkaeroja samassa työssä. Sama työ määritellään niin tiukasti kuin se käytettävään aineiston puitteissa on mahdollista eli lähes 300 muuttujan avulla. Samalla johdetaan uusi tapa ottaa regressioanalyysissä huomioon toimihenkilöiden palkanmuodostuksen kannalta keskeiset työn vaativuusluokitukset. Satoja indikaattorimuuttujia vaativa työn piirre tiivistetään yhteen työn vaativuutta kuvaavaan jatkuvaan muuttujaan, joka on helposti tulkittavissa ja käytettävissä. Tutkimuksessa osoitetaan myös, että ratkaisu ei vaikuta lainkaan palkkadiskriminaatiosta tehtäviin johtopäätöksiin.

Tutkimuksen tulosten mukaan teollisuussektorin toimihenkilöillä esiintyy palkkaeroja samaa työtä tekevien miesten ja naisten välillä. Suuruudeltaan tämä palkkadiskriminaatio on tyypillisessä toimihenkilötehtävässä noin 8 logaritmista prosenttia (noin 8 tavallista prosenttia). Loppuosa eli noin kaksi kolmannesta sukupuolten keskipalkkojen välisestä erosta johtuu miesten ja naisten erilaisista henkilökohtaisista ominaisuuksista ja eroista heidän työnsä piirteissä. Arvio palkkadiskriminaatiosta on aiempia palkkaerotutkimuksia selvästi alempi.

Yksittäisen toimihenkilön kohtaama palkkasyrjintä vaihtelee melko paljon, ja yli 5000 henkilöllä se on negatiivinen. Kahdella ja puolella tuhannella toimihenkilöllä palkkadiskriminaatio taas ylittää 20 log-prosenttia. Erityisen suurta syrjintää on ylempien toimihenkilöiden keskuudessa ja paperi- ja graafisessa teollisuudessa. Poikkeuksellisen pientä se taas on teknisten toimihenkilöiden parissa ja talonrakennus- ja meijerialalla. Koulutustaso ei vaikuta millään säännönmukaisella tavalla palkkadiskriminaatioon. Tutkimuksessa esitellään myös uusi tapa havainnollistaa palkkasyrjintää. Sen mukaan naisten tulisi hankkia noin kuusi vuotta miehiä enemmän työkokemusta, jotta he muuten samanlaisessa työssä saisivat samaa palkkaa. Vaihtoehtoisesti naisten tulisi hankkia yhden keskiasteen tutkimuksen verran lisäkoulutusta, jotta syrjivä palkkaero kuroutuisi umpeen.

Palkkadiskriminaatio riippuu yli 20 tekijästä, joista sukupuolten erilainen ikäkohtelu on selvästi voimakkain yksittäinen syrjintää aiheuttava palkanmuodostuksen piirre. Miehet ja naiset aloittavat työuransa likimain samasta tilanteesta,

mutta iän tuoma työkokemus kasvattaa miesten palkkaa huomattavasti nopeammin kuin naisten. Ero kumuloituu siten, että eläkeiän kynnyksellä diskriminaatio on melkein 14 log-%.

Toimihenkilön kohtaama palkkasyrjintä keskimäärin kasvaa, jos hänen työyksikkönsä on selvästi ympäristöään naisvaltaisempi, ja toimihenkilön palkan poikkeama keskipalkasta vaikuttaa diskriminaatioon samalla tavalla. Palkkasyrjintä on alhaisinta mediaanipalkan lähellä olevissa tuloluokissa ja suurinta hyvin matalasti tai korkeasti palkatuissa tehtävissä. Tutkimuksessa saadaan myös epäsuoraa näyttöä siitä, että työnantajan vaihto muodostaa ylimääräisen riskin naisten palkkakehitykselle.

Sisällys

1 Johdanto	1
1.1 Palkkaerojen problematiikkaa	1
1.2 Palkkadiskriminaation käsite	1
1.3 Aikaisempi palkkaerotutkimus, työn tavoitteet ja sisältö	3
2 Palkkadiskriminaatiotutkimuksen käsitteitä ja menetelmiä	5
2.1 Palkkadiskriminaatiofunktio	5
2.2 Oaxaca-hajotelma	7
2.3 Varovaisuusperiaate ja kovarianssianalyysi	11
2.4 Segregaation ja heterogeenisuuden vuorovaikutuksesta	12
3 Aineiston ja muuttujien esittely	15
3.1 Aineisto	15
3.2 Muuttujat ja niiden tunnuslukuja	16
3.3 Joitakin tilastollisia näkökohtia	26
4 Palkkamallin estimointi	30
4.1 Jatkuvien muuttujien muunnosten valinta	30
4.2 Laaja palkkamalli LPM	32
4.3 Mallin LPM supistaminen malliksi PM1	33
4.4 Toimihenkilöiden palkanmuodostuksen arviointia	37
5 Palkkadiskriminaation tarkastelu	41
5.1 Empiirinen palkkadiskriminaatiofunktio	41
5.2 Palkkasyrjintä teollisuuden eri osissa	43
5.3 Palkkadiskriminaatio ja jatkuvat muuttujat	46
5.4 Naisvaltaisuuden merkitys	50
5.5 Keillä palkkasyrjintä on erityisen suurta tai pientä?	51
5.6 Miesten ja naisten erilaisesta kohtelusta	52
6 Tulosten luotettavuuden ja oikeellisuuden arviointi	54
6.1 Palkkamalli GPM1 ja palkkadiskriminaation luottamusväli	54
6.2 Mahdollisten virhespesifikaatioiden etsiminen malleilla PM2 - PM15	54
6.3 Yhteenveto luvusta 6	58

7 Johtopäätökset	61
Lähteet	64
Liite	67

1 Johdanto

1.1 Palkkaerojen problematiikkaa

Miehet saavat keskimäärin naisia korkeampaa palkkaa käytännössä kaikilla työmarkkinoilla. Suomessa miesten keskipalkka on jopa muutamia kymmeniä prosentteja naisten keskipalkkaa suurempi riippuen siitä, mitä talouden sektoria tarkastellaan. Vaikka naisten osallistumisella työelämään on maassamme pitkät perinteet, ei Suomi sukupuolten palkkaeron suhteen juurikaan poikkea muista teollisuusmaista. (Korkeamäki – Kyyrä 2002, 3.) Tämä on herättänyt vilkasta yhteiskunnallista keskustelua palkkaeron syistä. Ilmiötä pidetään yleisesti korjaamista vaativana epäkohtana, ja se nousee säännönmukaisesti esiin mm. tulopoliittisissa neuvotteluissa. Se oli myös yksi kevään 2003 eduskuntavaalikamppailun teemoista (Pohjanpalo 2003a ja 2003b).

Ajoittain julkisessa keskustelussa esitetään, että sukupuolten välinen palkkaero olisi kokonaisuudessaan seurausta naisiin työmarkkinoilla kohdistuvasta diskriminaatiosta. Tällainen näkökulma on monessa mielessä pulmallinen. Miesten ja naisten välinen palkkaero on monimutkaisen prosessin tulos, ja täysin tasarvoisestikin toimivilla työmarkkinoilla voi esiintyä palkkaeroja jommankumman sukupuolen hyväksi. Toisaalta sukupuolten keskipalkkojen välisen eron umpeen kurominen ei välttämättä riitä poistamaan syrjintää kaikilta talouden sektoreilta. Näkökulma on ongelmallinen myös siksi, että voidaan määritellä useita sukupuoleen perustuvan syrjinnän lajeja ja kysyä, mitä palkkaeroa kasvatavaa prosessia diskriminaatiolla tarkoitetaan. Esimerkiksi naisten pääsyä paremmin palkattuihin perinteisesti miehille kuuluviin tehtäviin voidaan vaikeuttaa tai naisille voidaan maksaa miehiä alemmaa palkkaa samasta työstä; ensimmäistä kutsutaan rekrytointi- ja jälkimmäistä palkkadiskriminaatioksi. Edelleen miehillä ja naisilla saattaa olla erilaiset mahdollisuudet hankkia itselleen työelämässä etenemiseen tarvittavia tietoja ja taitoja tai miesten tekemä työ saatetaan arvioida naisten tekemää työtä arvokkaammaksi. Lisäksi miehet työskentelevät tyypillisesti naisia korkeapalkkaisemmissa tehtävissä. Työmarkkinoilla voi siis esiintyä useita sukupuolten väliseen palkkaeroon vaikuttavia mekanismeja, ja niistäkin vain osa on luonteeltaan syrjiviä. Ne ovat erillisiä mutta samanaikaisia ilmiöitä, ja ne menevät helposti sekaisin keskenään. Niiden erottaminen toisistaan on kuitenkin tärkeää, sillä eri mekanismeja tutkitaan eri tavoin. Myös saatujen tulosten luotettavuus riippuu keskeisesti tutkittavan ilmiön onnistuneesta rajaamisesta.

1.2 Palkkadiskriminaation käsite

Tämä työ käsittelee miesten ja naisten välistä palkkadiskriminaatiota Suomen teollisuussektorin toimihenkilöillä vuonna 2000. Esitellään aluksi palkkadiskriminaation käsite.

Palkkasyrjinnän tutkiminen perustuu inhimillisen pääoman teoriaan. Koska mm. Asplund (1993a) käsittelee aihetta kattavasti, esitellään tässä vain teorian keskeiset piirteet. Sen mukaan työntekijä osaa tehtävänsä sitä paremmin, mitä pidempi muodollinen koulutus tai työura hänellä on takanaan. Oman tehtävän hoitamiseen tarvittavat tiedot ja taidot kehittyvät myös, jos tehtävässä toimitaan kauan. Nämä kolme tekijää kasvattavat työntekijän inhimillistä osaamispääomaa. Mitä enemmän hän on kerännyt sitä, sen tehokkaammin hän työnsä tekee ja sen suurempaa palkkaa hänelle voidaan maksaa. Palkkaerot ovat siis seurausta henkilöiden välisistä tuottavuuseroista, joiden taas katsotaan johtuvan heidän työntekoon olennaisesti liittyvien ominaisuuksiensa erilaisuudesta. Vaikkapa korkeammin koulutetulle ja siksi työnsä tehokkaammin tekevälle kuuluu siitä isompi korvaus. Itse työhön liittyy myös ominaisuuksia, jotka vaikuttavat olennaisesti siitä maksettavaan palkkaan. Esimerkkinä mainittakoon työn vaativuustaso: haastavampaan työhön liittyy yleensä suurempi palkkio. Tietyn työntekijän palkka riippuu siis hänen ominaisuuksistaan ja hänen tekemänsä työn luonteesta. Kutsuttakoon näitä työntekijän palkan määrittäviä seikkoja hänen taustaominaisuuksikseen: osa niistä kuvaa työntekijän henkilökohtaisia ominaisuuksia, osa hänen tekemänsä työn luonnetta. Määritellään edelleen, että kaksi henkilöä tekevät samaa työtä, kun heidän taustaominaisuutensa ovat samat sekä henkilökohtaisten että työhön liittyvien seikkojen osalta. Tällöin voidaan arvioida, että he tekevät yhtä tuottavaa työtä.

Palkkadiskriminaatioksi kutsutaan sitä, että samaa työtä tekevä mies ja nainen saavat eri palkkaa. Keskeistä määrittelyssä on juuri saman työn käsite. Näin määritellyn palkkadiskriminaation esiintyminen on nimittäin yhtäpitävää sen kanssa, että yhtä tehokkaasta työstä ja siis samanlaisista taustaominaisuuksista palkitseminen vaihtelee sukupuolen mukaan. Sukupuolten välisistä (työhön tai henkilöön liittyvien) taustaominaisuuksien eroista syntyvää palkkaeroa ei sen sijaan pidetä palkkadiskriminaationa. (Vartia – Kurjenoja 1992, 2.) Tutkijat ovat erimielisiä siitä, kuinka sama työ tulisi määritellä, koska ei ole lainkaan yhdentekevää, mitä samalla työllä tarkoitetaan. Sen määritelmässä otetaan suoraan kantaa siihen, mitä osaa palkkaerosta pidetään muusta kuin palkkasyrjinnästä johtuvana. Jos sama työ määritellään tiukasti useaa kriteeriä käyttäen, voivat miehet ja naiset myös olla useammalla taustaominaisuudella mitattuna erilaisia. Koska näistä poikkeavuuksista syntynyttä palkkaeroa ei tulkita palkkadiskriminaatioksi, ymmärretään se tällöin suppeasti. Vastaavasti saman työn väljä määrittely johtaa laveaan näkemykseen ja yleensä korkeampaan arvioon palkkasyrjinnästä.

Tässä tutkielmassa sama työ määritellään niin tiukasti kuin se käytettävissä olevan aineiston perusteella on mahdollista. Näin menetellään, jotta voitaisiin olla varmoja siitä, että palkanmuodostukseen vaikuttavat seikat on otettu riittävällä tarkkuudella huomioon ja että arviot palkkadiskriminaatiosta koskevat nimenomaan yhtä tuottavaa työtä tekeviä henkilöitä. Toisaalta miehet ja naiset työskentelevät palkkatasoiltaan erilaisissa tehtävissä ja yrityksissä, ja tämä

segregaationa tunnettu ilmiö kasvattaa merkittävästi sukupuolten välistä palkkaeroa. Tutkielman palkkadiskriminaatiokäsite on haluttu puhdistaa myös siitä, koska se on erillinen ja oman tutkimuksensa ansaitseva prosessi. Esimerkiksi tehtävien erilaisista palkkatasoista johtuvaa palkkaeroa ei tässä tutkielmassa pidetä palkkadiskriminaationa, vaan katsotaan, että toimenkuvan vaihtuessa saman työn määritelmä ei enää täyty. (Vrt. Koev 1996, 1–4.)

Myöhemmin tässä tutkielmassa jaetaan miesten ja naisten välinen kokonaispalkkaero palkkadiskriminaatiosta ja kaikesta muusta johtuvaan osaan. Tällä ei tarkoiteta, että kaikki palkkadiskriminaation ulkopuolelle jäävä olisi kokonaisuudessaan jotenkin ongelmatonta tai hyväksyttävää palkkaeroa. Esimerkiksi miesten valikoituminen naisia vaativampiin tehtäviin saattaa toki heijastella alussa mainittua rekrytointidiskriminaatiota; se ei vain kuulu tämän tutkielman piiriin. (Koev 1996, 1–3.) Pikemminkin jaottelussa on kyse huomion keskittämisestä yhteen tiettyyn asiaan, jotta juuri siitä voitaisiin tehdä mahdollisimman tarkka analyysi. Empiirisessä tutkimuksessa tämä on erityisen tärkeää, koska tutkittavaa ilmiötä kuvaamaan on rakennettava tilastollinen malli, jonka perusteella päättely tehdään. Mitä selkeämmin aihe on rajattu, sen paremmin malli kyetään konstruoimaan ja sen tarkempaa päättelyä voidaan tehdä. Aiheen tarkasta rajauksesta riippuu myös, koskeeko päättely juuri kiinnostuksen kohteeksi valittua ilmiötä vai onko se sekaantunut johonkin toiseen samalta näyttävään prosessiin. Nämä ovat yleisiä empiirisen tutkimuksen haasteita, ja myöhemmin pohditaan, miten ne tulisi ottaa huomioon juuri palkkadiskriminaation tapauksessa.

1.3 Aikaisempi palkkaerotutkimus, työn tavoitteet ja sisältö

Suomessa on tehty viimeisen kymmenen vuoden aikana lukuisia tutkimuksia miesten ja naisten palkkaeroista. Mm. Asplund (1996), Vartiainen (1998, 1999 ja 2001) ja Lilja (1997) ovat tutkimuksissaan arvioineet taustaominaisuuksilla selittymättömän¹ osan olevan vajaan kymmenen ja kahdenkymmenen prosentin välillä. Vartia ja Kurjenoja (1992), Pasanen (1995) ja Koev (1996) ovat tutkineet palkkaeroa Suomen teollisuuden toimihenkilöiden keskuudessa nimenomaan palkkadiskriminaation näkökulmasta. Tämän tutkielman kannalta mielenkiintoisen vertailukohdan muodostavat näissä tutkimuksissa saadut tulokset noin kymmenen prosentin palkkadiskriminaatiosta. Arvioiden voimakas vaihtelu näyttäisi selittyvän pääasiallisesti juuri sillä, että sama työ ymmärretään tutkimuksissa eri tavoin. Yleensä palkkadiskriminaatio arvioidaan sitä pienemmäksi, mitä tiukemmin se määritellään.

¹ Tutkimuksissa tarkoitetaan taustaominaisuuksilla selittymättömällä osalla sitä osaa palkkaerosta, joka aiheutuu miesten ja naisten palkkayhtälöiden välisistä eroista. Tässä tutkielmassa sukupuolten palkkayhtälöiden eroista syntyvään palkkaeron osaan viitataan termillä palkkadiskriminaatio tai palkkasyrjintä. Jos sekaantumisen vaaraa ei ole, puhutaan lyhyesti vain diskriminaatiosta tai syrjinnästä.

Koska tutkittava ilmiö ja tarkasteltava talouden sektori ovat samoja kuin Vartiolla ja Kurjenojalla (mt.) ja Koevilla (mt.), ovat lukujen 1 ja 2 tutkimusongelman määrittely ja pohdinta joiltain osin yhteneviä heidän esityksiensä kanssa. Työ liittyy heidän tutkimuksiensa muodostamaan jatkumoon; tavoitteena on syventää tarkastelua ja tuoda uusia näkökulmia esiin. Vastausta haetaan mm. seuraaviin kysymyksiin:

- Esiintyykö teollisuussektorin toimihenkilöillä edellä määriteltyä palkkadiskriminaatiota?
- Kuinka suuri palkkadiskriminaatio on ja mistä tekijöistä se riippuu?
- Miten palkkadiskriminaatio vaihtelee teollisuussektorin eri osissa? (Vartia 1995, 15 ja Koev 1996, 4.)

Tämä luku johdattaa palkkadiskriminaation käsitteeseen, ja seuraavassa luvussa esitellään palkkasyrjinnän tutkimuksen menetelmä. Käytetty aineisto ja muuttujat esitellään luvussa 3, ja samalla johdetaan uusi tapa ottaa huomioon toimihenkilöiden palkanmuodostuksen kannalta keskeiset työn vaativuusluokitukset. Luvun lopussa pohditaan tilastollisia näkökohtia ja esitetään, miten alun perin aikasarja-analyysiä varten kehitettyä Frisch-Waugh-Lovell –lausetta voidaan hyödyntää monimutkaisten regressiotulosten tulkinnassa. Luvussa 4 estimoidaan yksityiskohtainen inhimillisen pääoman teoriaan perustuva tilastollinen palkkamalli, jonka perusteella luvussa 5 johdetaan syrjintää kuvaava palkkadiskriminaatiofunktio. Sen avulla arvioidaan palkkasyrjinnän suuruutta teollisuuden eri osissa ja sitä, miten palkkadiskriminaatio kehittyy toimihenkilön ominaisuuksien tai työn piirteiden muuttuessa. Luvun 5 lopussa havainnollistetaan miesten ja naisten erilaista kohtelua työmarkkinoilla vaihtamalla lähestymistapa tavanomaiseen verrattuna käänteiseksi. Kysytään, kuinka paljon osaavampia tai kokeneempia naisten tulisi olla, jotta he saisivat samaa palkkaa kuin miehet. Luvussa 6 pohditaan segregaatiota (eriytymistä) ja tutkitaan sen mahdollisesti aiheuttamia ongelmia, kun palkkadiskriminaatio arvioidaan selvästi heterogeenisestä aineistosta. Lisäksi estimoidaan arvioidulle palkkasyrjinnälle luottamusväli. Luvussa 7 tehdään yhteenveto saaduista tuloksista.

2 Palkkadiskriminaatiotutkimuksen käsitteitä ja menetelmiä

Palkkadiskriminaatiota olisi suhteellisen helppo tutkia, jos samanlaisia miehiä ja naisia voitaisiin kokeellisen tutkimuksen menetelmin siirrellä tehtävästä toiseen ja havainnoida, miten samasta työstä saatu palkka vaihtelisi sen mukaan, kumman sukupuolen edustaja sitä tekisi. Näin ei kuitenkaan voida monista eettisistä ja käytännön ongelmista johtuen tehdä, vaan päättely perustuu epäkokeellisesti syntyneeseen aineistoon. (Vartiainen 2001, 7.) Tällöin tuntuisi luontevalta tutkia syrjintää siten, että etsittäisiin mahdollisimman monelle aineiston naiselle vertailukohdaksi samaa työtä tekevä mies. Mikäli tällaisessa parittaisessa vertailussa havaittaisiin palkkaeroja, voitaisiin ne tulkita diskriminaatioksi – saivathan samaa työtä tehneet mies ja nainen eri palkkaa. Tällaisia suoraan vertailukelpoisia mies–nais-pareja on kuitenkin hyvin vähän, ja on kyseenalaista tehdä isoa perusjoukkoa koskevia päätelmiä pienen selkeästi valikoituneen osajoukon perusteella. Jos sama työ määritellään hyvin tiukasti, ei vertailukelpoisia pareja välttämättä löydy aineistosta lainkaan. Näistä piirteistä johtuen palkkadiskriminaation suuruus joudutaan arvioimaan teoreettisin keinoin. Kuvailtaan seuraavaksi, miten tämä tapahtuu.

2.1 Palkkadiskriminaatiodifunktio

Mitataan kohdassa 1.2 mainittuja työntekijän palkkaa määrittäviä taustaominaisuuksia k :lla muuttujalla x_1, x_2, \dots, x_k , joiden muodostamaa vektoria merkitään $\mathbf{x} = (x_1, x_2, \dots, x_k)'$. Nyt jokin tietty vektorin \mathbf{x} saama arvo kuvailee tietyn työtilanteen, ja jos se saa saman arvon kahdella eri henkilöllä, tekevät he samaa työtä. Tässä on syytä huomata, että vektori \mathbf{x} koostuu sekä henkilöä että työtä kuvailevista muuttujista. Työn samuuteen ei siis riitä, että työtehtävää kuvaavat tekijät ovat yhteneviä, vaan myös henkilökohtaisia ominaisuuksia mittaavien muuttujien on saatava samat arvot. Vaatimus palautuu aiempaan pohdintaan: henkilökohtaisilta ominaisuuksiltaan poikkeavien työntekijöiden katsotaan tekevän tuottavuudeltaan erilaista työtä. Olkoon edelleen

$$(1) \quad P_M = P_M(\mathbf{x}) = P_M((x_1, x_2, \dots, x_k)')$$

miehille ja

$$(2) \quad P_F = P_F(\mathbf{x}) = P_F((x_1, x_2, \dots, x_k)')$$

naisille estimoitu palkkayhtälö². Ne ovat tilastollisia kuvauksia siitä, kuinka odotettu palkka riippuu taustaominaisuuksista \mathbf{x} tai kuinka sukupuolia niistä pal-

² Kaavoissa viitataan luettavuuden parantamiseksi naisia koskevaan tietoon kirjaimella F.

kitaan. Palkkayhtälöiden avulla määritellään palkkadiskriminaatiofunktio (tuonempana lyhyemmin PDF)

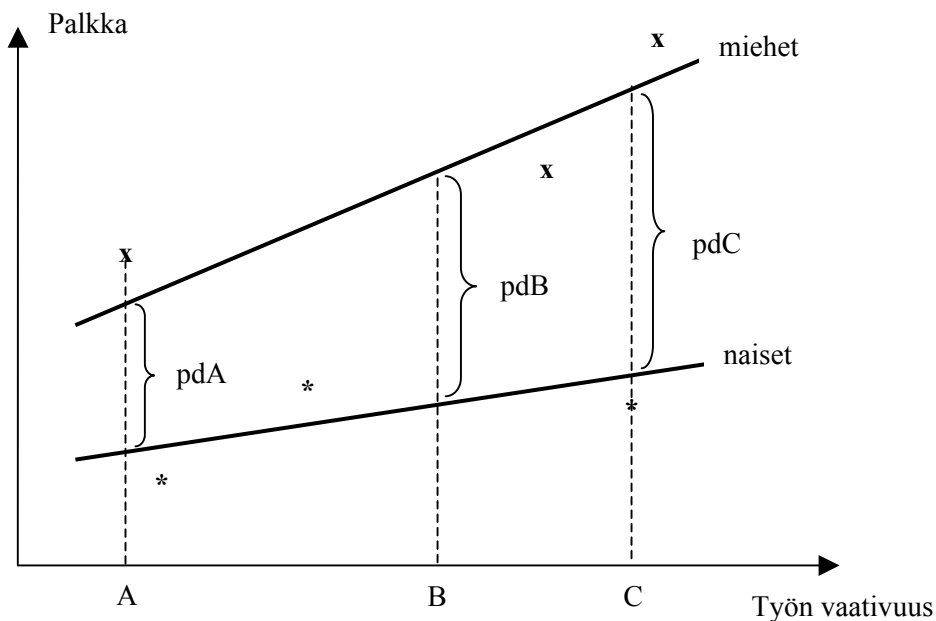
$$(3) \quad PDF(\mathbf{x}) \equiv P_M(\mathbf{x}) - P_F(\mathbf{x}),$$

joka mittaa sukupuolten välistä palkkaeroa *samassa* vektorin \mathbf{x} kuvailemassa työtilanteessa. (Vrt. Vartia 1995, 17 ja Koev 1996, 16.) Kun palautetaan mieleen kohdan 1.2 määritelmä, huomataan, että funktio (3) mittaa juuri saman työtilanteen \mathbf{x} kohdalla esiintyvää palkkadiskriminaatiota. Mikäli sukupuolia kohdellaan yhtäläisesti eli palkitaan samoin perustein, ei palkkasyrjintää esiinny. Palkkayhtälöt ovat identtisiä, ja $PDF(\mathbf{x})$ saa koko määrittelyjoukossaan arvon nolla. Jos taas $P_M(\mathbf{x})$ ja $P_F(\mathbf{x})$ ovat erilaisia argumenttinsa \mathbf{x} funktioita, poikkeaa $PDF(\mathbf{x})$ nolasta ja sen saama arvo riippuu siitä, mitä vektorin \mathbf{x} kuvaamaa työtilannetta tarkastellaan.

Kuten edellä todettiin, ei tiettyyn työtilanteeseen liittyvää palkkadiskriminaatiota voida havaita suoraan kokeellisesti, vaan se joudutaan korvaamaan teoreettisella vastineellaan. Siksi ainakin toinen funktion (3) termeistä on hypoteettinen. Jos esimerkiksi mies työskentelee vektorin $\mathbf{x} = \hat{\mathbf{x}}$ kuvaamassa toimessa, on $P_M(\hat{\mathbf{x}})$ palkka, jota hänen taustaominaisuuksillaan varustettu henkilö tyypillisesti saa. $P_F(\hat{\mathbf{x}})$ kuvaa sitä, kuinka paljon palkkaa hän saisi tästä samasta työstä, jos häntä palkittaisiin hänen taustaominaisuuksistaan $\hat{\mathbf{x}}$ naisten palkkayhtälön mukaan. Näiden erotus $PDF(\hat{\mathbf{x}})$ on positiivinen, jos työtilanteessa $\hat{\mathbf{x}}$ esiintyy palkkadiskriminaatiota miesten hyväksi. Tällöin $PDF(\hat{\mathbf{x}})$ ilmaisee, kuinka paljon kyseisen miehen palkka laskisi, jos häntä palkittaisiin (kohdeltaisiin) kuin naista. Jos taas tilanteessa $\mathbf{x} = \hat{\mathbf{x}}$ työskentelee nainen, vastaa palkkadiskriminaatiofunktio kysymykseen, kuinka paljon hänen palkkansa nousisi, jos häntä palkittaisiin (kohdeltaisiin) kuin miestä. PDF:n arvon määräytyminen ei siis riipu siitä, kumman sukupuolen edustaja tarkasteltavaa työtä tekee. (Koev 1996, 16.)

Palkkadiskriminaatiofunktion käsitettä havainnollistetaan kuviossa 1, jossa on oletettu, että palkka määräytyy ainoastaan työn vaativuuden perusteella. Miespuolista havaintoyksikköä on merkitty ristillä ja naispuolista tähdellä, ja kuuden havainnon aineiston perusteella on estimoitu sukupuolten palkkayhtälöt. Työtilannetta määrittää vain yksi tekijä eli työn vaativuus. Kuviosta 1 ei löydy suoraan vertailukelpoisia mies–nais-pareja, ja siinä mielessä sen hypoteettinen aineisto vastaa tosielämässä useimmiten esiintyvää tilannetta. Vaativuustasolla A työskentelee mies, ja hänen työtilanteessaan palkkasyrjinnäksi arvioidaan pystysuora etäisyys pdA . Vaativuustasolla C taas työskentelee nainen, ja hänen arvioidaan joutuvan etäisyyden pdC suuruisen palkkasyrjinnän kohteeksi. Yksikään aineiston kuudesta henkilöstä ei työskentele pisteen B määrittämässä työtilanteessa. Siinä esiintyväksi palkkadiskriminaatioksi arvioidaan etäisyys pdB .

Kuvio 1. Palkkadikriminointin tilastollinen arviointi epäkokeellisessa tilanteessa.



Tarkastelua voidaan siis laajentaa aineistossa toteutuneiden työtilanteiden ulkopuolelle kuvion 1 havainnollistamalla tavalla. Sukupuolten keskipalkkojen eron kannalta erityisen hyödyllisiä ovat arviot palkkadikriminointista miehille tai naisille tyypillisissä työtilanteissa. Tätä varten muodostetaan miesten keskimääräistä työtilannetta kuvaava vektori $\bar{\mathbf{x}}_M = (\bar{x}_1^M, \bar{x}_2^M, \dots, \bar{x}_k^M)'$. Siinä jokaiselle työtä kuvaavalle muuttujalle on annettu arvoksi sen keskiarvo miesten osajoukossa. Vastaavasti muodostetaan naisten tyypillistä työtilannetta mittaava vektori $\bar{\mathbf{x}}_F = (\bar{x}_1^F, \bar{x}_2^F, \dots, \bar{x}_k^F)'$. Niiden erotus $\bar{\mathbf{x}}_M - \bar{\mathbf{x}}_F$ kuvaa miesten ja naisten keskimääräisten työtilanteiden välistä eroa. Määritellyt vektorit ovat keskeisessä asemassa seuraavaksi esiteltävän Oaxaca-hajotelman johtamisessa.

2.2 Oaxaca-hajotelma

Yleensä palkan approksimoidaan riippuvan muuttujista x_1, x_2, \dots, x_k lineaarisesti ja miesten ja naisten palkkayhtälöt (1) ja (2) estimoidaan aineistosta regressioanalyysillä (ks. esim. Davidson – McKinnon 1993). Tällöin työntekijän i palkka on

$$(4) \quad P_i = P_M(\mathbf{x}_i) + e_i = \mathbf{b}'_M \mathbf{x}_i + e_i,$$

jos hän on mies, ja

$$(5) \quad P_i = P_F(\mathbf{x}_i) + e_i = \mathbf{b}'_F \mathbf{x}_i + e_i,$$

jos hän on nainen. \mathbf{b}_M ja \mathbf{b}_F ovat miesten ja naisten palkkayhtälöiden estimoitujen kertoimien muodostamat vektorit. Termi e_i on virhe, joka syntyy, kun työntekijän i palkkaa selitetään lineaarisesti hänen työtilannettaan kuvaavilla muuttujilla $\mathbf{x}_i = (x_{1i}, x_{2i}, \dots, x_{ki})'$. Se johtuu muuttujiin liittyvistä mittausvirheistä ja siitä, että ei käytetä oikeaa – tuntematonta – funktiomuotoa vaan sen lineaarista approksimaatiota. Käytännössä ei voida myöskään etukäteen tietää, mitkä muuttujat ovat relevantteja palkkayhtälön argumentteja: selittäjiksi saatetaan laittaa tarpeettomia muuttujia, tai jotkin keskeiset muuttujat saattavat jäädä pois tarkasteluista. Näin voi käydä mm. siksi, että monista palkkaan vaikuttavista tekijöistä ei ole saatavilla tietoja, tai niitä ei voida luotettavasti mitata. Viimeksi mainittuihin lukeutuvat monet työntekijän persoonallisuutta ja perhetaustaa kuvaavat seikat. Virhetermin käyttöä voidaan perustella vielä sillä, että palkanmuodostusta ei välttämättä kyetä kuvaamaan jonkin *funktion* avulla. Funktion määritelmään kuuluu olennaisesti yksikäsitteisyys: lähtöjoukon piste kuvautuu aina yhdeksi maalijoukon pisteeksi. Voidaan kuitenkin perustellusti vaatia, että kaksi samaa sukupuolta olevaa täysin samaa työtä tekevää henkilöä voivat saada *eri* palkkaa. (Voidaan vaikkapa ajatella, että jokin täysin ennustamaton satunnaismuuttuja on aiheuttanut eron.) Mikäli halutaan sallia tämä piirre palkanmuodostuksessa, ei sitä voida edes teoreettisesti kuvata funktiolla. Funktion ja jonkin satunnaismuuttujan summa on luontevampi lähtökohta.

Sukupuolten keskipalkkojen väliselle erolle voidaan nyt konstruoida esiteltyjä käsitteitä hyödyntäen luonteva ja palkkaerotarkasteluissa usein käytetty hajotelma. Koska kerroinvektorit \mathbf{b}_M ja \mathbf{b}_F estimoidaan siten, että virhetermien e_i keskiarvo on nolla, ovat sukupuolten keskipalkat

$$(6) \quad \bar{P}_M = P_M(\bar{\mathbf{x}}_M) = \mathbf{b}'_M \bar{\mathbf{x}}_M$$

ja

$$(7) \quad \bar{P}_F = P_F(\bar{\mathbf{x}}_F) = \mathbf{b}'_F \bar{\mathbf{x}}_F.$$

Niiden välinen erotus voidaan kirjoittaa muotoon

$$\begin{aligned} \bar{P}_M - \bar{P}_F &= P_M(\bar{\mathbf{x}}_M) - P_F(\bar{\mathbf{x}}_F) \\ &= P_M(\bar{\mathbf{x}}_M) - P_F(\bar{\mathbf{x}}_M) + P_F(\bar{\mathbf{x}}_M) - P_F(\bar{\mathbf{x}}_F) \\ (8) \quad &= \mathbf{b}'_M \bar{\mathbf{x}}_M - \mathbf{b}'_F \bar{\mathbf{x}}_M + \mathbf{b}'_F \bar{\mathbf{x}}_M - \mathbf{b}'_F \bar{\mathbf{x}}_F \\ &= \bar{\mathbf{x}}'_M (\mathbf{b}_M - \mathbf{b}_F) + \mathbf{b}'_F (\bar{\mathbf{x}}_M - \bar{\mathbf{x}}_F) \\ &= PDF(\bar{\mathbf{x}}_M) + \mathbf{b}'_F (\bar{\mathbf{x}}_M - \bar{\mathbf{x}}_F). \end{aligned}$$

Tätä kutsutaan Oaxaca-hajotelmaksi. (Oaxaca 1973.) Viimeisen rivin ensimmäinen termi on arvioitu palkkadiskriminaatio miesten tyypillisessä työtilanteessa

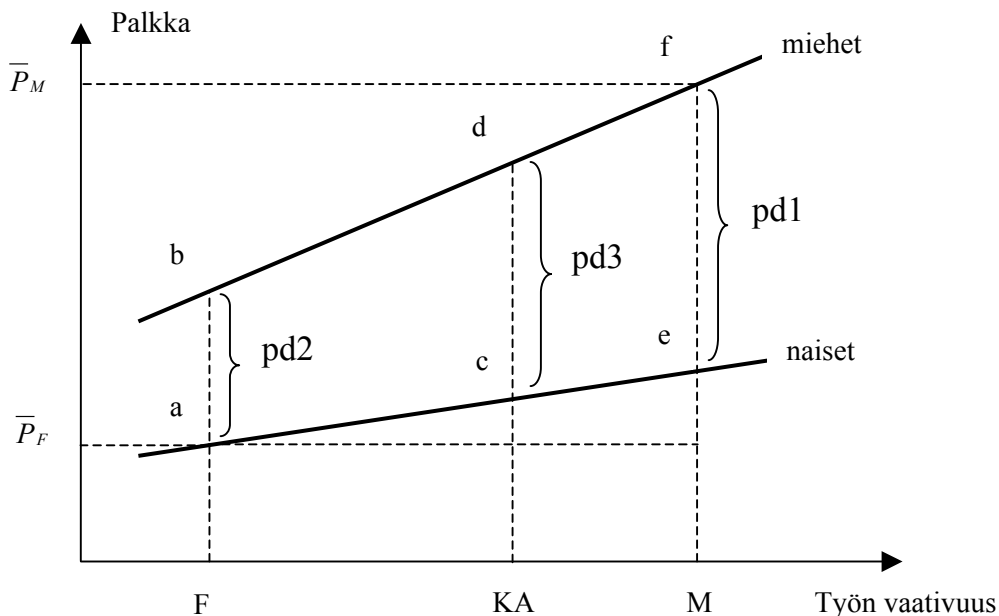
$\bar{\mathbf{x}}_M$. Toinen termi kuvaa sukupuolten keskimääräisten työtilanteiden poikkeavuudesta syntyvää keskipalkkojen eroa. Hajotelma ei ole yksikäsitteinen, vaan se voidaan kirjoittaa myös muodossa

$$\begin{aligned}
 \bar{P}_M - \bar{P}_F &= P_M(\bar{\mathbf{x}}_M) - P_F(\bar{\mathbf{x}}_F) \\
 (9) \quad &= \bar{\mathbf{x}}'_F (\mathbf{b}_M - \mathbf{b}_F) + \mathbf{b}'_M (\bar{\mathbf{x}}_M - \bar{\mathbf{x}}_F) \\
 &= PDF(\bar{\mathbf{x}}_F) + \mathbf{b}'_M (\bar{\mathbf{x}}_M - \bar{\mathbf{x}}_F).
 \end{aligned}$$

Tällöin palkkadiskriminaatiofunktion arvo lasketaan naisten keskimääräisen työtehtävän $\bar{\mathbf{x}}_M$ kohdalla.

Oaxaca-hajotelma on helppo tulkita graafisesti, jos oletetaan, että palkka riippuu vain yhdestä tekijästä – vaikkapa työn vaativuudesta. Kuvion 2 kuvitteellisessa tilanteessa miehet ovat naisia vaativammissa tehtävissä, koska työn vaativuutta kuvaavan (jatkuvan) muuttujan mieskeskiarvo M on korkeampi kuin naisten vastaava keskiarvo F. Lisäksi naisten palkka kasvaa hitaammin vaativuustason nousun myötä. Koko aineistosta laskettu vaativuuden keskiarvo on pisteessä KA, ja se on hiukan lähempänä miesten keskiarvoa, koska on oletettu, että aineisto on miesvaltainen.

Kuvio 2. Keskipalkkojen välisen eron jakautuminen eri pisteissä.



Jos keskipalkkojen erotusta $\bar{P}_M - \bar{P}_F$ tarkastellaan tyypillisen miestehtävän M kohdalla, saadaan palkkadiskriminaatioksi pystysuora siirtymä pd1. Loppuosaa palkkaerosta tulkitaan keskimääräisten vaativuustasojen erilaisuudesta johtuvaksi, ja sen merkitys arvioidaan naisten palkkayhtälön mukaan siirtymänä a - e. Tämä jaottelu vastaa hajotelmaa (8). Jos taas keskipalkkojen eroa tarkastellaan

pisteessä F eli naisten keskimääräisen työtehtävän kohdalla, saadaan palkkadiskriminaatioksi siirtymä pd2. Nyt vaativuustasojen erilaisuuden osuus palkkaerosta arvioidaan miesten palkkayhtälön mukaan ja tätä vastaa siirtymä b - f. Jaottelu vastaa hajotelmaa (9). Arvioitu palkkadiskriminaatio miesten keskiarvopisteessä kuvaa sitä, kuinka paljon miesten keskipalkka laskisi, jos heitä kohdeltaisiin kuin naisia. Naisten keskiarvopisteessä laskettu palkkadiskriminaatio taas mittaa sitä, kuinka paljon naisten keskipalkka nousisi, jos heitä kohdeltaisiin kuin miehiä. Ei voida yksikäsitteisesti sanoa, kumpi tarkastelutavoista on oikea; yleensä yhtälössä (8) esitetty tapa jakaa palkkaero kahteen osaan antaa hajotelmaa (9) suuremman arvion palkkadiskriminaatiosta. (Vartiainen 2001, 16.) Koevin mukaan molemmat arviot kannattaa raportoida (1996, 19). Niistä voidaan myös laskea miesten ja naisten osuuksilla painotettu keskiarvo, joka on aina suuruudeltaan niiden välissä. Sillä on mielenkiintoinen yhteys palkkadiskriminaatioon koko aineistosta lasketun keskimääräisen työtilanteen $\bar{\mathbf{x}} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k)'$ kohdalla, joten tarkastellaan sitä lähemmin.

Olkoot S_M miesten ja S_F naisten osuus aineistossa. Tällöin $S_M + S_F = 1$ ja $\bar{\mathbf{x}} = S_M \bar{\mathbf{x}}_M + S_F \bar{\mathbf{x}}_F$. Kerrotaan hajotelman (9) keskimäinen rivi naisten osuudella S_F ja hajotelman (8) toiseksi alin rivi miesten osuudella S_M . Kun lasketaan nämä kaksi tuloa yhteen, saadaan keskipalkkojen erolle seuraava hajotelma:

$$\begin{aligned}
 & S_F(\bar{P}_M - \bar{P}_F) + S_M(\bar{P}_M - \bar{P}_F) \\
 & = \bar{P}_M - \bar{P}_F \\
 & = S_F \bar{\mathbf{x}}'_F (\mathbf{b}_M - \mathbf{b}_F) + S_M \bar{\mathbf{x}}'_M (\mathbf{b}_M - \mathbf{b}_F) + S_F \mathbf{b}'_M (\bar{\mathbf{x}}_M - \bar{\mathbf{x}}_F) + S_M \mathbf{b}'_F (\bar{\mathbf{x}}_M - \bar{\mathbf{x}}_F) \\
 (10) \quad & = (S_F \bar{\mathbf{x}}'_F + S_M \bar{\mathbf{x}}'_M)(\mathbf{b}_M - \mathbf{b}_F) + \mathbf{b}'_F (S_M \bar{\mathbf{x}}_M - S_M \bar{\mathbf{x}}_F) + \mathbf{b}'_M (S_F \bar{\mathbf{x}}_M - S_F \bar{\mathbf{x}}_F) \\
 & = \bar{\mathbf{x}}' (\mathbf{b}_M - \mathbf{b}_F) + \mathbf{b}'_F (S_M \bar{\mathbf{x}}_M - (1 - S_F) \bar{\mathbf{x}}_F) + \mathbf{b}'_M ((1 - S_M) \bar{\mathbf{x}}_M - S_F \bar{\mathbf{x}}_F) \\
 & = PDF(\bar{\mathbf{x}}) + \mathbf{b}'_F (S_M \bar{\mathbf{x}}_M + S_F \bar{\mathbf{x}}_F - \bar{\mathbf{x}}_F) + \mathbf{b}'_M (\bar{\mathbf{x}}_M - S_M \bar{\mathbf{x}}_M - S_F \bar{\mathbf{x}}_F) \\
 & = PDF(\bar{\mathbf{x}}) + \mathbf{b}'_F (\bar{\mathbf{x}} - \bar{\mathbf{x}}_F) + \mathbf{b}'_M (\bar{\mathbf{x}}_M - \bar{\mathbf{x}}).
 \end{aligned}$$

Kokonaispalkkaero jakaantuu nyt kolmeen osaan. Hajotelman (10) erotus $\bar{\mathbf{x}} - \bar{\mathbf{x}}_F$ toisessa termissä kuvaa naisten tyypillisen työtilanteen poikkeamaa keskimääräisestä työtilanteesta $\bar{\mathbf{x}}$. Kuviossa 2 tämä on erotus KA - F. Sen osuus palkkaerosta arvioidaan naisten palkkayhtälön mukaan, ja sitä vastaa kuviossa 2 siirtymä a - c. Ensimmäinen termi on puolestaan palkkadiskriminaatio keskimääräisessä työtilanteessa, ja kuviossa 2 se on pystysuora siirtymä pd3. Kolmas termi on miesten keskimääräisen työtilanteen poikkeaman $\bar{\mathbf{x}}_M - \bar{\mathbf{x}}$ osuus palkkaerosta arvioituna miesten palkkayhtälön mukaan. Kuviossa 2 sitä kuvaa siirtymä d - f. Lisäksi havaitaan, että yhtälöiden (8) ja (9) palkkadiskriminaatioista laskettu painotettu keskiarvo on täsmälleen palkkadiskriminaatio keskimääräisen työtilanteen $\bar{\mathbf{x}}$ kohdalla.

Kuvion 2 mukaan tulkittuna keskipalkkojen ero a - f voidaan siis jakaa osiin kolmea eri reittiä a - e - f, a - b - f tai a - c - d - f pitkin, ja vastaavasti palkkadiskriminaatiolle voidaan antaa kolme eri arviota pd1, pd2 tai pd3. Yleensä käytetään vain kahta ensimmäistä, koska molemmat pyrkivät vastaamaan luonteviin kysymyksiin, ja niille löytyy siksi selkeä tulkinta. Ei kuitenkaan ole olemassa yksiselitteistä kriteeriä arvioida, kumpi niistä on oikea, joten myös kolmas lasikutapa puoltaa paikkaansa. Se on eräänlainen kompromissi kahdesta ensimmäisestä, ja sen etuna on, että työtilanteiden erilaisuudesta aiheutuvan palkkaeron arviointiin käytetään molempien sukupuolten palkkayhtälöitä. Edun merkitys korostuu, jos ne poikkeavat huomattavasti toisistaan. Tällöin hajotelmat (8) ja (9) voivat antaa hyvin erilaisia arvioita palkkadiskriminaatiosta.

2.3 Varovaisuusperiaate ja kovarianssianalyysi

Tutkielman nollahypoteesi on, että palkkadiskriminaatiota ei esiinny. Jos saadaan riittävän selvää näyttöä sukupuolten erilaisesta palkanmuodostuksesta, astuu vaihtoehtoinen hypoteesi voimaan. Kun lähtökohta on asetettu näin, ei voida estimoida miesten ja naisten palkkayhtälöitä (1) ja (2) erikseen ja käyttää suoraan niiden erotusta palkkadiskriminaatiofunktiona (3). Se kävisi päinsä, jos voitaisiin olla varmoja, että erikseen estimoidut yhtälöt kuvaavat palkanmuodostusta täysin oikein. Kohdassa 2.1 luetelluista syistä on kuitenkin selvää, että yhtälöt ovat monin tavoin puutteellisia ja osin jopa virheellisiä approksimaatioita todellisesta tilanteesta. Tämä ei suinkaan tarkoita, että ne olisivat käyttökelvottomia tilastolliseen päättelyyn, vaan että niiden perusteella voidaan tehdä johtopäätöksiä palkkadiskriminaatiosta vain, jos näyttö on riittävän selkeää. Tätä kutsutaan varovaisuusperiaatteeksi (Vartia 1995, 18). Vaaditaan, että saadun näytön voimakkuuden on ylitettävä sen hankkimiseen liittyvä epävarmuus, mikä on yleinen tilastollisen päättelyn lähtökohta. Palkkadiskriminaatiofunktioon (3) hyväksytään vain ne muuttujat, joiden estimoidut kertoimet miesten ja naisten palkkayhtälöissä poikkeavat tilastollisesti merkitsevästi toisistaan. Muiden muuttujien kohdalla katsotaan, ettei niistä palkitseminen ole erilaista sukupuolten välillä. Niiden kerroin palkkadiskriminaatiofunktiossa on nolla. Jos yksikään kerroin ei todistettavasti ole erilainen miesten ja naisten palkkayhtälöissä, ei aineisto anna aiheutta johtopäätökseen palkkadiskriminaatiosta. (Vartia 1995, 17.) Tällöin nollahypoteesi sukupuolten samanlaisesta palkanmuodostuksesta jää voimaan.

Varovaisuusperiaatteen huomioiva menetelmä on estimoida miehille ja naisille yhteinen palkkayhtälö kovarianssianalyysillä (Wonnacott – Wonnacott 1979, 105–107). Se on regressio- ja varianssianalyysin yhdistelmä, ja sitä on luontevaa käyttää, kun selittäjinä on sekä luokittelevia että jatkuvia muuttujia. Palkanmuodostuksen vapaa vaihtelu sukupuolen mukaan sallitaan lisäämällä selittäjiksi naisindikaattorin n interaktiota muiden muuttujien kanssa. Teknisesti interaktiot muodostetaan siten, että ne saavat alkuperäisen muuttujan arvon, jos kysessä on nainen ja muutoin arvon nolla. Esimerkiksi vaativuuden erilainen vaikutus suku-

puolen mukaan voidaan sallia käyttämällä selittäjänä vaativuuden lisäksi interaktiota $n \cdot \text{vaativuus}$. Olkoon nyt vaativuuden estimoitu kerroin 5 ja interaktion -1 . Tällöin miesten palkkayhtälössä vaativuuden kerroin on 5 ja naisten palkkayhtälössä $5 - 1 = 4$. Interaktion kerroin ilmaisee suoraan miesten ja naisten yhtälöiden kertoimien erotuksen, ja jos se poikkeaa tilastollisesti merkitsevästi nolasta, esiintyy kyseisen muuttujan kohdalla palkkadiskriminaatiota. Kovarianssianalyysin etu on juuri se, että jokaisesta muuttujasta voidaan testata tilastollisesti, onko sen kerroin erilainen sukupuolten välillä. Edelleen voidaan määritellä korkeamman asteen interaktioita, joilla mahdollistetaan mallin kertoimien vaihtelu aineiston eri ositteissa. Jos esimerkiksi muuttuja ylt on ylempien toimihenkilöiden indikaattori, sallitaan interaktiolla $n \cdot ylt \cdot \text{vaativuus}$ ylemmille naistoimihenkilöille vaativuuden poikkeava kerroin. (Vrt. Koev 1996, 9–14.) Kovarianssianalyysin käyttö myös helpottaa palkkadiskriminaation analysointia, koska tilastollisesti merkityksettömien interaktioiden poiston vuoksi palkkadiskriminaatiofunktioon jää alkuperäistä palkkayhtälöä huomattavasti vähemmän argumentteja (Vartia 1995, 17).

2.4 Segregaation ja heterogeenisuuden vuorovaikutuksesta

Sama malli kykenee harvoin kuvaamaan palkanmuodostusta riittävän tarkasti koko aineistossa, koska palkkayhtälön kertoimet vaihtelevat sen eri ositteissa. Havainnollistetaan aluksi kahden esimerkin avulla millaisiin virhepäätelmiin voidaan ajautua, jos tätä heterogeenisuutta ei oteta huomioon. Niissä käytetään termiä aineisto melko väljästi.

Tarkastellaan kahdesta sektorista muodostuvaa taloutta, jossa palkka P määräytyy lineaarisesti ja deterministisesti ainoastaan työntekijän iän I perusteella. Sektorilla 1 tämä riippuvuus on $P = \alpha + \beta_1 I$ ja sektorilla 2 $P = \alpha + \beta_2 I$. β_1 on suurempi kuin β_2 eli sektorilla 1 palkitaan iästä paremmin suuremman palkan muodossa. Tämä on varsin luonteva oletus, jos ajatellaan, että iän mukanaan tuoma elämäkokemus on sektorin 1 työssä tarpeellisempaa kuin sektorin 2 työssä. Jälkimmäinen voi olla vaikkapa jokin rutiininomainen helposti opittava työ. Oletetaan edelleen, että sektorilla 1 työskentelee 90 miestä ja 10 naista ja sektorilla 2 90 naista ja 10 miestä. Lisäksi sukupuolet ovat sekä sektorien sisällä että koko aineistossa verrattuna keskimäärin saman ikäisiä. Nyt esimerkin taloudessa esiintyy palkkaero miesten hyväksi. Jos lasketaan sektoreista piittaamatta aggregaattitason Oaxaca-hajotelma (9), supistuu sen jälkimmäinen keskimääräisten työtilanteiden eroa mittaava osa pois, koska mitään eroa ei yksinkertaisesti ole. Ainoa työtilannetta määrittävä tekijä on ikä, jonka keskiarvo on sukupuolilla sama. Palkkaero tulkitaan kokonaisuudessaan ikäkohtelun erosta eli palkkadiskriminaatiosta johtuvaksi. Päätelmä on kyllä siinä mielessä oikea, että suurin osa naisista on todellakin sektorilla 2 huonomman ikäkohtelun piirissä. Sen sijaan tulkinta palkkadiskriminaatiosta on kyseenalainen, koska kummankin sektorin sisällä sukupuolia kohdellaan tasa-arvoisesti. (Vrt. Vartia ym. 1992, 1–7.) Virhe-

päätelmä johtuu siitä, että aineistoa käsiteltiin homogeenisena, vaikka sitä olisi pitänyt tarkastella sektoreittain. Aggregaattitason Oaxaca-hajotelma ei siis riitä edes näin yksinkertaisessa tilanteessa palkkadiskriminaation kuvaukseksi. Tämä ei suinkaan ole itse hajotelman vaan aineistoon sopimattoman liian jäykän mallin vika.

Jatketaan edellisen esimerkin hengessä. Oletetaan nyt, että molemmilla sektoreilla on 50 miestä ja naista ja että sukupuolten ikäkeskiarvot ovat äsken kuvattulla tavalla samat. Olkoot sektoreilla erilaiset palkkayhtälöt miehille ja naisille seuraavasti: $P_{1M} = 4I$, $P_{1F} = 2I$, $P_{2M} = 2I$ ja $P_{2F} = 4I$ (vakio on nolla). Keskipalkkojen ero on tällöin nolla. Jos nyt lasketaan sektorijakoa huomioimaton aggregaattitason hajotelma (9), supistuu myös ensimmäinen palkkadiskriminaatiota mittaava termi pois, koska molemmille sukupuolille arvioidaan ikäkertoimeksi kolme. Näin käy siitä huolimatta, että sektorilla 1 palkkadiskriminoidaan naisia ja sektorilla 2 miehiä. Jälleen virhepäätelmä johtuu aineistoon sopimattomasta liian jäykästä mallista. Samalla huomataan, kuinka pelkkien keskipalkkojen vertailu antaa riittämättömän kuvan syrjinnästä.

Esimerkit valaisevat yhtä aihepiirin keskeistä haastetta. Empiirisen palkkadiskriminaatiotutkimuksen tavoitteeksi ei riitä pelkkä erilaisten kertoimien löytäminen. Yhtä tärkeää on kyetä osoittamaan luotettavasti ne osat taloudesta, joissa *ei* esiinny palkkadiskriminaatiota eli kertoimet ovat samoja sukupuolten välillä. Ensimmäisessä esimerkissä epäonnistuttiin jälkimmäisessä tavoitteessa ja toisessa esimerkissä ensin mainitussa. Molempien tavoitteiden samanaikaista saavuttamista vaikeuttavat (ainakin) seuraavat seikat:

1. Palkkayhtälöiden kertoimet voivat olla erilaisia eri osissa taloutta, vaikka palkkadiskriminaatiota ei esiintyisi lainkaan.
2. Miehet ja naiset ovat segregoituneet (eriytyneet) erilaisten kertoimien piiriin.
3. Etukäteen ei tiedetä, miten talous tulisi jakaa kohdassa 1 mainittuihin osiin.

Viimeinen kohta on olennainen lisäys esimerkkeihin, joissa sektorijako oletettiin tunnetuksi. Palkkayhtälöä estimoitaessa joudutaan aina jossain määrin arvaamaan, missä talouden osissa palkanmuodostus voisi poiketa päättelyä vaarantavalla tavalla. Käytännössä kovarianssianalyyseissä pitää interaktioiden käytöllä varmistaa, että malli on riittävän joustava. Näin mahdollistetaan kertoimien muuttuminen aineiston eri ositteissa. Aina voidaan tietysti vaatia aineiston hienompaa ositusta ja erilaisia kertoimia entistä pienemmille sektoreille. Jos tällainen disaggregointi viedään äärimmäisyyksiin, jokainen työntekijä saa palkkaa omilla kertoimillaan, eikä mitään palkkadiskriminaatiota voida löytää. Tällöin sukupuolten palkkaero johtuu kokonaisuudessaan eriytymisestä ja kaikkien katsotaan tekevän eri työtä, mikä on tulkintana vailla mieltä. Mallinnusteknisesti ollaan syyllistytty tilastolliseen ylihdollistamiseen (engl. overconditioning), ja ainoa mallin parametrien suhteen tyhjentävä tunnusluku on aineisto itse. Frek-

ventistisiä todennäköisyystarkasteluja ei tällöin voida lainkaan tehdä. (Ekholm 2000, 104–109.) Ei voida yksikäsitteisesti sanoa, mikä on oikea aggregoinnin taso; ainakin osittain se palautuu tutkimuksessa käytetyn saman työn ja sitä kautta palkkadiskriminaation määritelmään. Jos ne ovat hyvin väljiä, ei välttämättä tuntuisi tarpeelliselta käyttää kovin joustavaa mallia. Rungas interaktioiden käyttö ja siis tarkka mallinnus ovat puolestaan luontevasti sopusoinnussa tiukkojen määritelmien kanssa.

Yhteenvedona voidaan todeta, että palkkadiskriminaatiotutkimuksen kannalta segregatio on aito ongelma, joka saattaa vääristää johtopäätöksiä. Sukupuolten segregoituminen talouden eri sektoreille toimii eräänlaisena vipusimena, joka voi saada aineiston heterogeenisuuden näyttämään aggregaattitason tarkasteluissa palkkadiskriminaatiolta. Jos vaikkapa miehet ovat päätyneet naisia useammin yrityksiin, joissa iästä yleisesti palkitaan paremmin, näkyy tämä aggregaattitasolla miesten naisia suurempana ikäkertoimena. Vaikka tähän valikoitumiseen liittyykin syrjiviä piirteitä, ei sitä voida tulkita palkkadiskriminaatioksi. Toisaalta sektorien sisäiset syrjintämekanismit saattavat jäädä havaitsematta, jos palkkayhtälö estimoidaan liian aggregoidulla tasolla. Näiden seikkojen merkitystä tutkitaan tarkemmin luvussa 6.

3 Aineiston ja muuttujien esittely

3.1 Aineisto

Palkkadiskriminaation empiirinen analyysi perustuu Teollisuuden ja Työnantajain Keskusliiton (TT) vuonna 2000 keräämään toimihenkilöitä koskevaan palkka-aineistoon. TT tekee vuosittain joulukuussa jäsenyrityksilleen palkkatiedustelun, joka koskee periaatteessa kaikkia TT:n jäsenyrityksissä kyseisen vuoden lopulla työskennelleitä henkilöitä yritysten omistajia ja aivan ylimpään johtoon kuuluvia lukuun ottamatta (TT 2001, 55). TT ja sen edeltäjä Suomen Työnantajain Keskusliitto ovat laatineet palkkatilastoja kymmenien vuosien ajan (Kettunen – Marjanen 1992, 1), joten tiedonkeruumenetelmät ovat vakiintuneita ja aineistoa voidaan pitää laadukkaana ja luotettavana. Koska se on lisäksi erittäin laaja, kannattaa sen sisältämää informaatiota hyödyntää analyysissä mahdollisimman paljon. Tutkielman mallinnustekniikka onkin kauttaaltaan erittäin aineistolähtöinen.

Analyysi rajattiin täysi-ikäisiin kokoaikaisesti työskenteleviin toimihenkilöihin, jotka olivat olleet aineiston keruuhetkellä voimassa olleessa työsuhteessaan vähintään kolme kuukautta (vrt. esim. Lilja 1995, 7). Harjoittelijat ja osa-aikaiset toimihenkilöt jätettiin pois tarkasteluista, koska on aihetta olettaa, että heidän palkanmuodostuksensa poikkeaa tavanomaisesta. Lisäksi muutamia selvästi teollisuussektoriin kuulumattomia tai muuttujien koodauksiltaan epäselvimpiä sopimusaloja poistettiin analyysin helpottamiseksi. Ensin mainittuun ryhmään kuuluvat esimerkiksi Yleisradio ja ilmaliikenne, jälkimmäiseen taas toimistotekninen kauppa ja graafisen teollisuuden toimituksellinen henkilökunta. Mukana olevat sopimusalat luetellaan liitteessä 1. Tiedoiltaan puutteellisimmat havainnot jouduttiin niinikään poistamaan, joskin niitä oli hyvin vähän aineiston kokoon nähden. Rajatussa aineistossa on 149 343 toimihenkilöä, joista miehiä on 63 ja naisia 37 prosenttia. Se muodostaa kokonaisuutensa analyysissä mukana olevista sopimusaloista, joten tarkasteluissa ei ole mitään otantaan liittyviä epävarmuustekijöitä. Koska TT määrittelee itsensä teollisuudessa ja sitä lähellä olevilla aloilla toimivien yritysten etujärjestöksi (TT 2003a), on analyysissä mukana muutamia vain väljästi teollisuuden piiriin luettavia tai sitä palvelevia aloja. Saatuja tuloksia voidaankin pitää TT:n itseään koskevan määritelmän hengessä ei vain teollisuuden, vaan laajemmin teollisuussektorin toimihenkilöitä koskevina. Toki pitää huomioida, että johtopäätökset koskevat täydellisellä tarkkuudella ainoastaan analyysissä mukana olevia aloja. Toisaalta niitä on karsittu melko varovaisesti, joten tulosten yleistettävyyden on tuskin kärsinyt.

Aineisto sisältää jokaisesta toimihenkilöstä hänen taustaansa, kuten koulutusta kuvaavia tietoja ja useita eri palkkalajeja. Koska TT laatii palkka-aineistonsa palvelemaan lähinnä työmarkkinatoimintaa (Kettunen ym. 1992, 1–2), se ei sel-

laisenaan sovellu palkkadiskriminaation tutkimiseen. Siksi siihen luotiin muutamia uusia muuttujia ja seuraavaksi esitellään miten ne johdettiin. Samalla esitellään niistä laskettuja tunnuslukuja, jotka havainnollistavat miesten ja naisten erilaisia työtilanteita.

3.2 Muuttujat ja niiden tunnuslukuja

Ei ole yhdentekevää kuinka analysoitava ansiokäsite valitaan, koska peruspalkan lisäksi maksettavien tulospalkkioiden ja lisien määräytyminen on monitahoinen prosessi. Mm. Vartiainen on pohtinut, kuinka ylitöiden kautta saadut tulot pitäisi huomioida (2001, 22–23). Selitettävää palkkamuuttujaa muodostettaessa valittiin lähtökohdaksi toimihenkilön säännöllisen työajan peruskuukausipalkka, koska se kuvaa hyvin, kuinka työntekijää palkitaan hänen henkilökohtaisista ominaisuuksistaan. Se ei pidä sisällään erilaisia ylityölisä, tulospalkkioita tai luontoisetuja (TT 2001, 56–57), joiden suuruus määräytynee pikemminkin yritys- ja toimialakohtaisten työntekokulttuurien perusteella kuin henkilötasolla. Lisäksi miehet tekevät tyypillisesti naisia enemmän ylitöitä, ja laajan ansiokäsitteen käyttö saattaisi johtaa palkkadiskriminaation yliarviointiin (Korkeamäki ym. 2002, 6–7). Toimihenkilön tuntipalkka laskettiin kuukausipalkasta säännöllisen viikottaisen työajan perusteella, ja selitettävä muuttuja on euroissa ilmaistun tuntipalkan logaritminen muunnos. Työajalla skaalaamisen ansiosta sitä ei tarvitse enää käyttää mallissa selittäjänä. Lisäksi sen varianssi on niin pieni, että sen merkitys selittäjänä olisi vähäinen. Aineistossa naisten keskituntipalkka on 11,61 euroa ja miesten 28,4 prosenttia korkeampi eli 14,91 euroa. Myöhemmin laskettavat hajotelmat koskevat logaritmisten keskipalkkojen eroa, joka on 0,25 eli 25 log-prosenttia (myöhemmin lyhyemmin log-%).

Selittävät muuttujat voidaan jakaa kolmeen ryhmään sen mukaan, mitä palkanmuodostuksen piirrettä ne mittaavat. Inhimillistä pääomaa kuvaaviin muuttujiin kuuluvat ikä ja voimassa olevan työsuhteen kesto, joka laskettiin toimihenkilön meneillään olevan työsuhteen alkuhetken ja aineiston keruuhetken välisenä erotuksena. Jos työsuhte oli keskeytynyt, laskettiin sen kestoksi työsuhteen alku- ja loppuhetken välinen aika. Toimihenkilön ikä kuvaa hänen koko työuransa aikana hankkimaa kokemusta ja työsuhteen kesto hänen pätevyyttään nykyisessä toimessaan. Kummankin mittayksikkö on vuosi. Muodollinen koulutus on kolmas inhimillisen pääoman komponentti. TT:n aineistosta löytyy tieto toimihenkilön suorittamasta tutkinnosta, joka muunnettiin koulutusvuosiksi taulukon 1 esittämällä tavalla. Keskimääräistä tutkinnon suorittamiseen kulunutta aikaa määritettäessä käytettiin apuna Asplundin (1993b, 13) esittämää koulutusavainta ja TT:n omaa arviota tutkinnon kestosta (TT 2001, 59).

Taulukko 1. Tutkinnon muuntaminen koulutusvuosiksi

<i>Tutkinnon taso</i>	Perus- koulu	Keski- aste	Alin korkea-aste	Alempi korkea- kouluaste	Ylempi korkea- kouluaste	Tutkija- koulutus
<i>Koulutusvuodet</i>	9	11	13	14	16	20

Koulutustason lisäksi tiedossa on ala, jolta toimihenkilö on tutkintonsa suorittanut. Koulutusvuosista palkitseminen saattaa vaihdella sen mukaan miltä alalta ne on hankittu, ja tuntuisi luontevalta, että teollisuudessa olisi eniten hyötyä nimenomaan teknillisluonteisten alojen koulutuksesta (vrt. Asplund 1993b, 72–73). Niihin laskettiin maa- ja metsätalous-, luonnontieteellinen ja tekniikan ala, ja tämän ryhmän indikaattoriksi muodostettiin muuttuja klt. Muut alat toimivat vertailuryhmänä.

Taulukko 2. Toimihenkilön inhimillistä pääomaa kuvaavien muuttujien keskiarvot

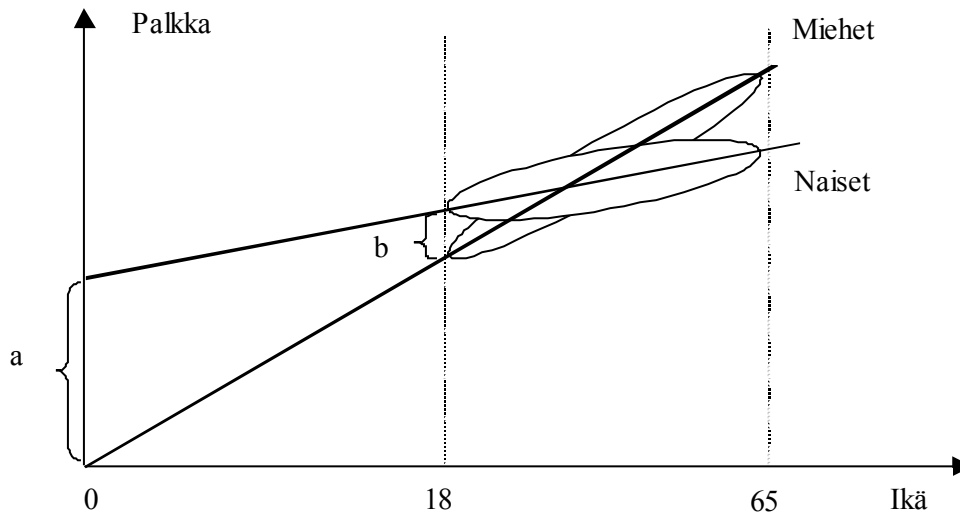
<i>Toimihenkilön ominaisuus</i>	<i>Miesten keskiarvo</i>	<i>Naisten keskiarvo</i>
Ikä	41,13	41,15
Työsuhteen kesto	11,85	12,39
Koulutusvuodet	12,81	12,06
Teknillisluonteinen koulutus	0,7030	0,2089

Taulukosta 2 huomataan, että sukupuolet ovat keskimäärin saman ikäisiä. Naisten muodolliseen koulutukseen tyypillisesti käyttämä aika on lukiota vastaavalla tasolla (12 vuotta), miesten taas jonkin verran sen yläpuolella. Sukupuolet eroavat jyrkästi toisistaan hankkimansa koulutuksen suuntautumisen suhteen. Vain viidennes naisista on saanut jonkin teknillisluonteisen alan koulutuksen, kun taas miehillä vastaava lukema on 70 prosenttia. Naiset ovat olleet nykyisessä työsuhteessaan miehiä kauemmin, mikä on sopusoinnussa pitkittäisaineistosta saatujen tulosten kanssa. Miehet näyttäisivät löytävän naisia helpommin hyvän uraputken ja saavan todennäköisemmin ylennyksiä (Lilja 1996, 27). Jos miesten nopeampaan urakehitykseen liittyy useita työnantajan vaihdoksia, näkyy se poikkileikkausaineistossa siten, että naiset viiptyvät työsuhteissaan miehiä pidempään.

Esiteltyjä inhimillistä pääomaa kuvaavia muuttujia voitaisiin käyttää (ja yleensä käytetäänkin) regressioanalyysissä sellaisenaan. Tällöin regressiovakion ja naisindikaattorin kertoimen tulokinnassa syntyisi kuitenkin hankaluuksia, joita kuvio 3 havainnollistaa. Siinä on oletettu, että palkka riippuu lineaarisesti ainoastaan iästä, jonka vaihteluväli on 18–65 vuotta. Vakio ilmaisee, kuinka paljon toimihenkilön palkaksi arvioidaan, kun kaikki selittäjät saavat arvon nolla. Naisindikaattorin kerroin a kuvaa miesten ja naisten palkkaeroa samassa tilanteessa. Koska yhdenkään toimihenkilön ikä ei ole nolla, ei vakiolle ja naisindikaattorin

kertoimelle saada mielekästä tulkintaa. Tilanne, jota ne kuvaisivat, on ekstrapoloitu liian kauas aineistosta.

Kuvio 3. Vakion ja naisindikaattorin kertoimen tulkinta



Kaikki tässä analyysissä käytettävät muuttujat skaalattiin siten, että vakio ja ennen kaikkea naisindikaattorin kerroin ovat järkevästi tulkittavissa. Jokaisen toimihenkilön iästä vähennettiin kahdeksantoista ja koulutusvuosista kahdeksan vuotta, ja regressioissa käytettiin näin saatuja muuttujia ikä18 ja kv8. Menettelyn syynä ovat aikaisemmassa tutkimuksessa saadut viitteet siitä, että alhaisimmilla palkkatasoilla diskriminaatio saattaisi olla negatiivinen (Vartiainen 2001, 76–87). Toisaalta juuri nuorimmat ja vähiten koulutetut saavat matalinta palkkaa. Muuttujien skaalauksella saadaan vakio ja naisindikaattorin kerroin kuvaamaan palkkausta juuri tässä aivan työelämänsä alussa olevien toimihenkilöiden ryhmässä. Jos naisilla olisi jonkinlainen alkupalkkaetu miehiin nähden, näkyisi se naisindikaattorin positiivisena kertoimena. Kuviossa 3 tällaista etua vastaa pystysuora etäisyys b.

Työehtosopimusten perusteella johdetut muuttujat muodostavat toisen tärkeän selittäjäjoukon. Toimihenkilöt jaotellaan vanhastaan ylempiin, teknisiin ja konttoritoimihenkilöihin³ sen perusteella, minkä työehtosopimuksen piiriin he kuuluvat (Vartiainen 2001, 30). Vastaavat toimihenkilöryhmien indikaattorimuuttujat ovat ylt, tet ja kot. Joillakin sopimusaloilla tekniset ja konttoritoimihenkilöt ovat saman alakohtaisen sopimuksen piirissä, eikä tällöin ollut tiedossa kumpaan kyseisistä toimihenkilöryhmistä työntekijä kuuluu. Tätä kahdesta ryhmästä yhdistettyä joukkoa kutsutaan vastaisuudessa muiksi toimihenkilöiksi, ja se muodostaa regressioanalyysin vertailuluokan. Teknisten, konttori-, ja muiden toimihenkilöi-

³ Konttoritoimihenkilöitä kutsutaan myös teollisuustoimihenkilöiksi.

den muodostamasta joukosta käytetään myöhemmin nimitystä alemmat toimihenkilöt erotukseksi ylemmistä toimihenkilöistä.

Taulukko 3. Miesten ja naisten jakautuminen toimihenkilöryhmiin

<i>Toimihenkilöryhmä</i>	<i>Ryhmän osuus aineistosta, %</i>	<i>Ryhmän osuus miehistä, %</i>	<i>Ryhmän osuus naisista, %</i>	<i>Naisia ryhmässä, %</i>
Konttorith:t	10,40	2,87	23,18	82,59
Tekniset th:t	12,75	16,99	5,56	16,14
Ylemmät th:t	37,21	46,06	22,16	22,06
Muut th:t	39,64	34,08	49,10	45,88

Taulukon 3 perusteella havaitaan, että ylemmät ja muut toimihenkilöt ovat suurimmat ryhmät. Tekniset toimihenkilöt ovat pääasiassa miehiä, kun taas konttoritoimihenkilöt enimmäkseen naisia; muiden toimihenkilöiden ryhmässä sukupuolijakauma on melko tasainen. Lisäksi sukupuolet ovat päätyneet melko erilaisiin toimihenkilöryhmiin. Miehet työskentelevät useimmin ylempänä toimihenkilönä ja naiset muuna toimihenkilönä. Harva nainen on päätenyt tekniseksi ja vielä harvempi mies konttoritoimihenkilöksi.

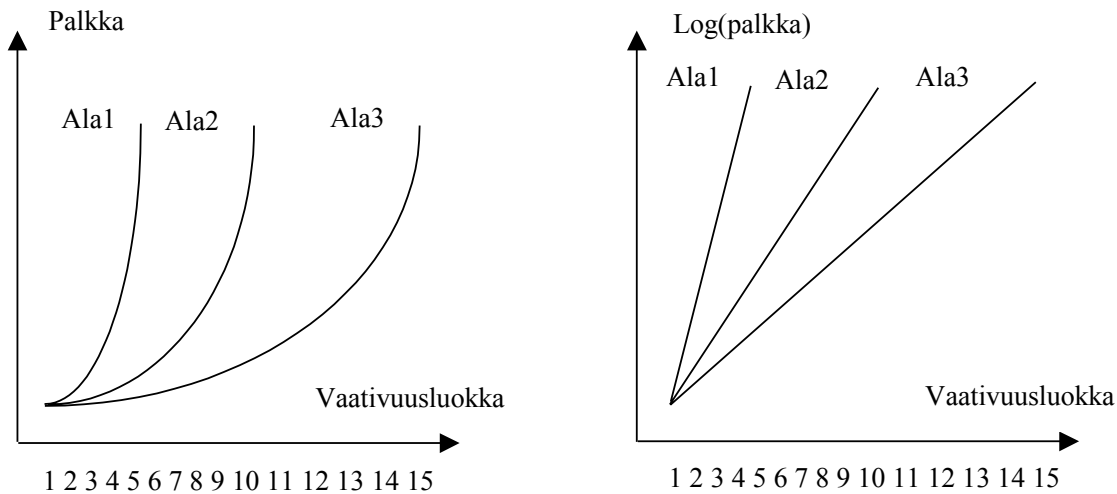
Toimihenkilöiden jaottelu ryhmiin on tärkeää, koska erilaisista työehtosopimuksista johtuen on aihetta odottaa, että palkanmuodostus vaihtelee ryhmittäin. Tähän viittaa myös aiempi tutkimus (Koev 1996, 39). Erityisesti alempien ja ylempien toimihenkilöiden työehtosopimukset poikkeavat selkeästi toisistaan. Ensin mainituissa käytetään työn vaativuusluokitusta, joka määrittelee palkan hyvinkin tarkkaan tehtävän vaatiman koulutuksen, työn tuoman vastuun ja sen rasittavuuden perusteella. Joissakin työehtosopimuksissa alemmat toimihenkilöt jaotellaan palkkaryhmiin. Koska erottelu perustuu samoihin edellä lueteltuihin kriteereihin, viitataan myös siihen vastaisuudessa termillä työn vaativuusluokitus. Ylemmillä toimihenkilöillä ei sen sijaan ole johdonmukaista vaativuusluokitusta, ja heitä koskevat sopimukset ovat huomattavasti väljempinä. Niissä kuvaillaan lähinnä yleisiä palkkauksessa hyviksi katsottuja periaatteita ja yrityksille annetaan melko vapaat kädet päättää ylempien toimihenkilöidensä palkitsemisesta. (Lilja 1996, 2 ja 38.)

Koska ylemmillä toimihenkilöillä ei ole vaativuusluokitusta, ei heille myöskään kyetä rakentamaan työn vaativuutta kontrolloivaa muuttujaa. Sen sijaan alemmille toimihenkilöille tämä voidaan tehdä. On kuitenkin ongelmallista, että työehtosopimukset ja niiden myötä vaativuusluokitukset ovat sopimusaloittaisia. Ne voivat vaihdella myös alan sisällä, mikä ilmenee liitteestä 1. Esimerkiksi teknisille ja konttoritoimihenkilöille on useilla aloilla erilliset sopimukset ja omat 6- ja 12-portaiset asteikot. Metalliteollisuuden asteikoissa taas on 9, 10 tai 15 portasta. Koska vaativuusluokkien lukumäärä vaihtelee sopimusaloittain ja osittain

alojen sisälläkin, eivät asteikot ole sellaisenaan vertailukelpoisia. Toisaalta ei ole järkevää luoda jokaisen asteikon jokaiselle vaativuusluokalle omaa indikaattorimuuttujaa. Tyypillisessä työn vaativuusasteikossa on kymmenen luokkaa ja aineistossa on 55 erilaista asteikkoa. Kun tämä moninaisuus otetaan huomioon, tarvittaisiin satoja indikaattorimuuttujia yhden ainoan työn piirteen kontrolloimiseen. Vaikka aineiston laajuus sinänsä mahdollistaisikin tällaisen menettelyn, olisi analyysissä välttämättömiä interaktioita erittäin vaikea muodostaa, ja tulosten tulkinnassa syntyisi melkoisia hankaluuksia.

Ongelmaa lähdettiin ratkomaan etsimällä Valtion Sääöstietopankki Finlexistä (Finlex 2003) eri alojen työehtosopimuksia, jotka olivat voimassa vuoden 2000 lopussa. Niistä löytyi yleensä palkkataulukot, joissa oli määritelty kuinka suurta palkkaa kuhunkin vaativuusluokkaan kuuluvalla toimihenkilölle pitää vähintään maksaa. Eri työehtosopimusten taulukoita tutkimalla havaittiin hyödyllinen ja tarkasteluja yksinkertaistava yhteys työn vaativuuden ja palkan välillä. Toimihenkilön siirtyminen häntä koskevalla asteikolla yhtä pykälää korkeampaan luokkaan johtaa yleensä hänelle kuuluvan palkan nousuun likimain vakioisella prosentilla. Tällöin palkka kasvaa eksponentiaalisesti siirryttäessä luokka kerrallaan vaativuusasteikon alimmalta portaalta ylimmälle. Kun siirrytään käyttämään selitettävänä muuttujana palkan logaritmia, syntyy sen ja toimihenkilön tekemän työn vaativuuden välille lineaarinen riippuvuus, mitä kuvio 4 havainnollistaa. Siinä on oletettu, että alalla 1 on käytössä 5-portainen, alalla 2 kymmenportainen ja alalla 3 15-portainen työn vaativuusluokitus. Alinta luokkaa on merkitty 1:llä. Vasemmanpuoleisessa osassa kuviota palkka kasvaa vakioisella alakohtaisella prosentilla siirryttäessä alemmasta yhtä ylempään vaativuusluokkaan. Oikeanpuoleisessa osassa on siirrytty tarkastelemaan palkan logaritmia, joka riippuu työn vaativuusluokasta lineaarisesti. Sen kerroin on sopimusalakohtainen: jotta likimain sama palkkavaihtelu saataisiin katettua eripituisilla asteikoilla, täytyy alalla 1 vaativuusluokan muutokseen liittyä suurempi palkankorotus kuin alalla 3.

Kuvio 4. Palkan riippuvuus vaativuusluokasta



Kaikilta aloilta ei enää tutkimuksen suoritusajankohtana (talvella 2002–2003) ollut saatavissa vuoden 2000 lopussa voimassa ollutta työehtosopimusta. Tämä ei liene suuri puute, koska alakohtaiset vaativuusluokitukset muuttuvat suhteellisen harvoin. Lisäksi eri aloilla käytettävät vaativuusasteikot (ks. liite 1) ilmenevät melkein aina suoraan aineistosta.

Ordinaaliasteikollista vaativuusluokitusta voidaan siis käyttää jatkuvana selittäjänä, mutta se ei vielä ratkaise koko ongelmaa. Koska luokitukset vaihtelevat aloittain ja lisäksi monilla sopimusaloilla on käytössä useampia asteikkoja, tarvittaisiin työn vaativuuden kontrollointiin edelleen 55 jatkuvaa muuttujaa. Kuvion 4 havainnollistamasta syystä kerroin on yleensä sitä pienempi, mitä enemmän asteikossa on luokkia. Tähän havaintoon perustuen johdettiin lineaarinen muunnos, joka yhdenmukaistaa vaativuusluokitukset. Koskeeseen jotakin tiettyä toimihenkilöä työehtosopimus, jossa vaativuusasteikko on $1, 2, \dots, z$ siten, että 1 on sen alin ja z korkein luokka. Olkoon edelleen $x \in \{1, 2, \dots, z\}$ kyseisen toimihenkilön tekemän työn vaativuus. Tällöin hänen tekemänsä työn yleisvaativuus on

$$(11) \quad y/va = (z-1)^{-1}(9x + z - 10).$$

Toimihenkilön työn vaativuutta alunperin kuvannut luku x on nyt skaalattu häntä koskevan vaativuusasteikon pituudella: moniportaiset asteikot kutistetaan ja lyhyet asteikot venytetään välille 1–10. Muunnos johdettiin siten, että alin luokka kuvautuu aina yhdeksi ja ylin kymmeneksi riippumatta alkuperäisen asteikon pituudesta. Loput lähtöasteikon luokat kuvautuvat desimaaliluvuiksi, ja ne sijaitsevat tasaisesti yhden ja kymmenen välillä. Kyseinen väli valittiin muunnoksen arvojoukoksi, koska tyypillisin teollisuuden toimihenkilöillä käytettävä vaativuusasteikko on juuri kymmenportainen. Olkoon työehtosopimuksessa esimer-

kiksi kuusiportainen vaativuusluokitus. Tällöin $z = 6$ ja muunnoskaavaksi (11) saadaan

$$(12) \quad y/va = (5)^{-1}(9x - 4).$$

Muuttujan *y/va* saamat arvot ovat nyt 1–2,8–4,6–6,4–8,2–10 vastaten kuutta alkuperäistä luokkaa.

Muunnos (11) kutistaa moniportaisia ja venyttää lyhyitä asteikkoja siten, että niiden kulmakertoimet muuttuvat likimain samoiksi. Kuviossa 4 alan 1 kerroin laskee, koska sen kohdalla vaaka-akselia venytettäisiin ja alan 3 kerroin kasvaisi, koska sen vaaka-akselia kutistettaisiin. Alalla 2 arvot pysyisivät ennallaan, koska siellä käytetään jo valmiiksi kymmenportaista asteikkoa. Skaalauksen (11) ansiosta ei tarvitse käyttää joka asteikolle omaa vaativuusmuuttujaa, vaan *y/va* voidaan käyttää selittäjänä sellaisenaan. Satoja indikaattorimuuttujia vaativa ordinaaliasteikollinen työn piirre saadaan tiivistettyä yhteen jatkuvaan muuttujaan. Lisäksi *y/va* on helposti tulkittavissa. Se ilmaisee yksinkertaisesti toimihenkilön sijainnin häntä koskevalla välillä 1–10 skaalatulla vaativuusasteikolla. Erilaiset vaativuusasteikot saadaan yhteismitallisiksi ja keskenään vertailukelpoiksi⁴. Aiemmasta tutkimuksesta ei kyetty löytämään vastaavaa kymmeniä asteikkoita yhdenmukaistavaa ratkaisua, joten mitä ilmeisimmin kyseessä on uusi tapa ottaa huomioon teollisuuden toimihenkilöiden monenkirjavat vaativuusluokitukset. Niitä on kyllä hyödynnetty muissa analyyseissä, mutta niitä ei ole yhdenmukaistettu tässä esitetyllä tavalla koko teollisuutta kattavaksi yhdeksi muuttujaksi (esim. Vartia ym. 1992, Koev 1996, Lilja 1995 ja 1997 tai Asplund 1993b). Muutamit tutkimukset tarkastelevat niin kapeaa sektoria, että vaativuusluokitukset ovat luonnostaan samanlaisia (esim. Vartiainen 1998 ja 1999). Aina ei työn vaativuutta ole käytetty selittäjänä ollenkaan (esim. Vartiainen 2001). Työehtosopimusten luonteesta johtuen se on kuitenkin tärkein yksittäinen toimihenkilön palkkaa määrittävä työn piirre. Se on otettava huomioon, jos halutaan saada tarkka kuva toimihenkilöiden palkanmuodostuksesta.

Yleisvaativuus *y/va* kyettiin laskemaan yhteensä 72 000 toimihenkilöille eli noin puolelle aineistosta. Lopuille toimihenkilöille annettiin *y/va* arvoksi nolla ja yleisvaativuustiedon puuttumiselle luotiin indikaattori *y/vapt*. Selkeästi suurin osa puuttuvista havainnoista koostuu ylemmistä toimihenkilöistä, joille ei siis ole olemassa vaativuusluokituksia lainkaan. 22 000 alemman toimihenkilön kohdalla *y/va* saa arvon nolla, koska alkuperäinen vaativuusluokka ei ollut tiedossa tai muuttujan koodaus jäi epäselväksi. Jos se sai vaikkapa arvon ”3” tai ”f”, oli tieto hyödytön, jos ei selvinnyt mihin asteikkoon se viittasi. Siksi liitteessä 1 on lueteltu vain ne asteikot, jotka kyettiin aineistoa ja työehtosopimuksia tutkimalla tunnistamaan ja joita pystytään analyyseissä hyödyntämään.

⁴ TT:n oman arvion mukaan vaativuusluokitusten vertailu sopimusalarajojen yli on mahdotonta, koska asteikot ovat erilaisia (TT 2003b, 13–14). *y/va* käyttö yhdistettynä regressioanalyysiin on siis tähän selvä parannus.

Muuttuja ylva luonnollisesti yksinkertaistaa todellista tilannetta. Palkan logaritmin riippuvuus vaativuusluokasta ei ole kaikissa asteikoissa täysin lineaarinen, eivätkä kertoimet muutu täysin samoiksi skaalauksella (11). Toisaalta lineaarinen yhteys havaittiin työehtosopimusten palkkataulukkoita tutkimalla, ja se on syytä varmentaa myös empiirisesti. Siksi otettiin lähempään tarkasteluun ne toimihenkilöt, joille ylva pystyttiin muodostamaan ja vertailtiin erilaisten heille muodostettujen palkkamallien selitystasetta, kun vaativuutta kontrolloitiin eri tavoin. Jos selitystasetta olisi romahtanut joko siirryttäessä luokittelevista jatkuviin muuttujiin tai niistä edelleen ylvaan, olisi tehtyä yksinkertaistusta voitu pitää kyseenalaisena. Koska alakohtaisten palkkatasojen huomioiminen on välttämätöntä ylvan kertoimen estimoinnissa, käytettiin jokaisessa mallissa selittäjinä sopimuslainsäädäntöä. Niiden ja vaativuusmuuttujien lisäksi ei käytetty muita selittäjiä, koska regressioilla haluttiin nimenomaan tutkia, löytyvätkö työehtosopimusten palkkataulukkoissa havaitut piirteet myös aineistosta. Sopimuslainsäädäntöä käytön vuoksi vertailukohdan muodosti regressio, jossa vain ne olivat selittäjinä. Sen selitystasetta oli 0,312. Kun malliin lisättiin jokaisen vaativuusasteikon jokaiselle portaalle (ks. liite 1) muodostetut 519 indikaattoria, nousi selitystasetta 42 prosenttiyksiköllä 0,734:ään. Tarkemmin ei työn vaativuutta voi käytettävissä olevan aineiston puitteissa kontrolloida. Seuraavaksi mallissa muutettiin vaativuusmuuttujat jatkuviksi, mutta sallittiin jokaiselle asteikolle oma kerroin. Selitystasetta laski 0,727:ään. Pudotus on hyvin pieni, kun otetaan huomioon, että vaativuuden kontrollointiin tarvittiin enää alle 60 muuttujaa eli murto-osa edellisessä mallissa käytetystä määrästä. Kuvion 4 havainnollistama asteikoittainen log-lineaarinen riippuvuus vaikuttaa aineistosta saadun tuloksen valossa huomattavan vahvalta. Lopuksi kaikille asteikoille tehtiin muunnos (11) ja työn vaativuuden kuvaamiseen käytettiin ainoastaan näin saatua ylva. Mallin selitystasetta oli 0,681 eli vaativuuskontrollin lisäyksen myötä saavutetusta selityskyvyn noususta jäi kahden yksinkertaistuksen jälkeen jäljelle 37 prosenttiyksikköä. Ne laskevat siis yhteensä palkkamallin selitystasetta viisi prosenttiyksikköä, mitä voidaan pitää pienenä lukuna saavutettuihin etuihin verrattuna. Erityisesti ensin tehty siirtyminen luokittelevasta jatkuvaan vaativuuden kontrollointiin vaikuttaa harmittomalta yksinkertaistukselta⁵. Asteikoiden yhdenmukaistaminen muunnoksella (11) johti jo suurempaan selitystasteen laskuun. Vaikka kyseinen 4,6 prosenttiyksikön pudotus on suhteellisen vähäinen, ei pelkkää mallin teknistä sopivuutta tutkimalla täysin saada selville, muuttaako toimenpide palkkasyrjinnästä tehtäviä johtopäätöksiä. Siksi sen merkitystä tutkitaan tarkemmin kohdassa 6.2. Siinä osoitetaan, että vaikka ylvan kertoimet hiukan vaihtelevatkin skaalauksen (11) jälkeen, ei niiden pakottaminen samoiksi muuta arviota palkkadiskriminaatiosta. Ylva johdettaessa tehdyt kaksi ratkaisua ovat siis tavoitteen kannalta harmittomia yksinkertaistuksia.

⁵Harmittoman yksinkertaistuksen korjaaminen todenmukaiseksi ei muuta johtopäätöksiä (Vartia 1988, 307).

Ennen ylvan käyttöä on vielä otettava kantaa siihen, miten toimihenkilön tekemän työn vaativuus määräytyy. Jos työnantaja saa arvioida sen vapaasti, on hänellä periaatteessa tilaisuus sijoittaa toisen sukupuolen edustajat järjestelmällisesti alempiin luokkiin. Tällöin työn vaativuutta käytettäisiin palkkadiskriminaation välineenä. Ei olisi järkevää käyttää sitä kuvaavia muuttujia selittäjinä, koska palkkasyrjintä aliarvioitaisiin. Tämän selvittämiseksi eri alojen työehtosopimuksista etsittiin tietoja toimihenkilön vaativuusluokan määräytymisestä (Finlex 2003). Havaittiin, että arvioinnin objektiivisuus on useimmiten pyritty tavalla tai toisella varmistamaan. Yleensä toimihenkilöiden edustajien osallistuminen vaativuusluokkien määrittelyyn on sallittu. Työehtosopimuksissa on myös lueteltu hyvin tarkasti useita työn arviointikriteerejä, kuten sen tuoma vastuu tai sen vaatima asiantuntemus. Niistä annetaan pisteitä, joiden summa määrittää toimihenkilön vaativuusluokan ja hänen vähimmäispalkkansa. Joillakin aloilla erityinen arviointiryhmä määrittelee tiettyyn vaativuusluokkaan kuulumisen kriteerit, tai vaativuusluokan määritelmässä on lueteltu eksplisiittisesti tehtäviä, jotka sen piiriin kuuluvat. Työnantajalla on siis harvoin tilaisuus yksinään päättää toimihenkilön tekemän työn vaativuudesta. Luokitukset sisältävät aitoa työtä kuvaavaa informaatiota ja niiden käyttö regressioissa on perusteltua.

Ylvan avulla voidaan tutkia, ovatko sukupuolet sijoittuneet vaativuustasoltaan erilaisiin tehtäviin. Keskiarvot jouduttiin laskemaan niistä toimihenkilöistä, joille ylva pystyttiin muodostamaan, ja siksi tulokset koskevat vain alempia toimihenkilöitä. Miesten keskimääräinen yleisvaativuus kyseisessä joukossa on 5,68 ja naisten 4,64. Ylvan vaihteluvälin 1–10 keskipiste taas on 5,5. Kun verrataan keskiarvoja siihen, havaitaan, että miehet ovat sijoittuneet tyypillisesti hiukan oman asteikkonsa keskikohdan yläpuolelle ja naiset selvästi sen alapuolelle. Keskiarvojen ero on melko tarkkaan yksi, mikä vastaa kymmenportaisella asteikolla juuri yhden ja 15-portaisella vaativuusasteikolla noin puolentoista luokan suuruista eroa.

Viimeinen työehtosopimuksista johdettu muuttuja on työskentelykunnan kalteusluokka. Työehtosopimuksissa kunnat jaetaan hintatasonsa mukaan kahteen ryhmään, ja elinkustannuksiltaan korkeammissa kunnissa työskentelystä maksetaan korkeampaa palkkaa. Muuttuja kalleulk on kalliin kunnan indikaattori.

Kolmanteen ja viimeiseen ryhmään kuuluvat selittäjät ovat luonteeltaan erilaisia palkkatasoja kontrolloivia apumuuttujia. Toimihenkilöt jakaantuvat 38 sopimus-alalle, jotka on lueteltu liitteessä 1. Jokaiselle tehtiin indikaattori ja regressioiden vertailuluokaksi valittiin teleliikenneala. Aineistossa on 1639 yritystä, ja jokaiselle yli 200 hengen yritykselle luotiin indikaattori. Niitä on 126 kappaletta, ja yhteensä niissä työskentelee 66 % aineiston toimihenkilöistä. Koska yhdellä pienimmistä sopimusaloista toimii vain yksi yritys, piti sen indikaattori poistaa selittäjien joukosta täydellisen multikollinearisuuden välttämiseksi. Se ja alle 200 hengen yritykset muodostavat vertailuluokan. Tehtävää kontrolloivia muuttujia konstruoitaessa käytettiin suoraan TT:n omaa tilastonimikkeistöä. Se jaottelee

toimihenkilöt toimiperheisiin, jotka edelleen jakautuvat 75 eri tehtävään. Esimerkiksi toimiperheeseen koulutus kuuluu kaksi tehtävää: koulutus- ja opetustyö ja työnopastustyö. (TT 2001, 60–61.) Joistakin toimihenkilöistä oli tiedossa tehtävämääntäjän puutteellisen arvon vuoksi ainoastaan toimiperhe, eikä varsinainen tehtävä ilmennyt aineistosta. Ensisijaisesti heidät pyrittiin kohdentamaan johonkin kyseisen toimiperheen tehtävään heidän saamansa palkan perusteella. Jos tätä ei ollut mahdollista tehdä yksiselitteisesti, jouduttiin kyseisen toimihenkilön kohdalla tyytymään ainoastaan toimiperhettä koskevaan tietoon. Näin syntyi TT:n alkuperäisten tehtäväluokkien lisäksi viisi uutta toimiperhettä osoittavaa luokkaa, joissa työskentelee yhteensä 9788 toimihenkilöä. Sitä voidaan pitää aineiston kokoon nähden suhteellisen pienenä lukuna. Tehtäväindikaattoreita on siis yhteensä 80 ja regressioiden vertailuluokaksi valittiin henkilöstöpalvelutyö.

Koska luokkia on paljon, olisi taulukon 3 kaltainen tarkastelu yritysten, sopimusalojen ja tehtävien kohdalla melko sekava. Sukupuolten erilaisesta sijoittumisesta ei saataisi selvää kuvaa. Paremmiin sitä havainnollistaa ns. Duncan-indeksi, joka tunnetaan myös dissimilaarisuusindeksinä. Oletetaan, että aineisto on jaettu jollain kriteerillä k :hon luokkaan ja olkoot m_i luokan i osuus miehistä ja vastavasti n_i luokan i osuus aineiston naisista. Duncan-indeksi määritellään

$$(13) \quad D = 0,5 \cdot \sum_{i=1}^k |m_i - n_i|.$$

Se on aina nollan ja yhden välillä, ja se ilmaisee kuinka suuren osan miehistä (tai vaihtoehtoisesti naisista) olisi siirryttävä luokasta toiseen, jotta sukupuolet olisivat jakautuneet niihin samalla tavalla. Se saa arvon yksi, jos eriytyminen on täydellistä eli miehet ja naiset ovat kokonaan omissa luokissaan. Jos taas indeksi saa arvon nolla, ovat sukupuolet jakautuneet niihin tasaisesti. (Vartiainen 2001, 20.) Taulukossa 4 raportoidaan Duncan-indeksin saamat arvot.

Taulukko 4. Duncan-indeksit sopimusaloittain, yrityksittäin ja tehtävittäin laskettuna

<i>Minkä jaottelun suhteen laskettu</i>	<i>Duncan-indeksi</i>
Sopimusalat	0,2221
Yli 200 hengen yritykset	0,1723
Tehtävät	0,5339

Viidenneksen miehistä tai naisista tulisi vaihtaa sopimusalaa, jotta sukupuolet olisivat jakautuneet niille samalla tavalla. Tilanne on lähes sama, jos segregoitumista tarkastellaan yrityksittäin. On tosin huomioitava, että alle 200 hengen yritykset muodostavat yhden ison luokan, mikä alentane arviota. Teollisuuden mies- ja naistoimihenkilöt ovat päätyneet selvästi erilaisiin ammatteihin: enemmän kuin joka toisen naisen tai miehen pitäisi vaihtaa tehtävää, jotta sukupuolet olisivat jakautuneet niihin tasaisesti.

Taulukossa 5 on tehty yhteenveto esitellyistä muuttujista. Keskeisten epälineaaristen riippuvuuksien huomioimiseksi analyysissä tullaan käyttämään iän ja työsuhteen keston neliöjuurimuunnoksia. Kohdassa 4.1 esitetään, miten niihin on päädytty.

Taulukko 5. Muuttujat

<i>Toimihenkilön ominaisuus tai työn piirre</i>	<i>Kontrolloiva muuttuja</i>	<i>Mahdolliset arvot ja mittayksikkö</i>
Sukupuoli	Naisindikaattori n	0 tai 1
Toimihenkilöryhmä	Ylempien, teknisten ja konttoritoimihenkilöiden indikaattorit ylt, tet ja kot	0 tai 1
Työn vaativuus	Yleisvaativuus ylva	1,0-10,0. Ei tietoa=0.
Työn vaativuustiedon puuttuminen	Puuttumisen indikaattori ylvapt	0 tai 1
Työpaikan sijaintikunnan kalteusluokka	Kalliin kunnan indikaattori kalleulk	0 tai 1
Ikä	Ikä18 (=ikä-18) ja sen neliöjuuri	0–47 vuotta
Nykyisen työsuhteen kesto	Tysk ja sen neliöjuuri	0,25 - vuotta
Koulutukseen käytetty aika	Kv8 (=koulutusvuodet-8)	1–12 vuotta
Koulutuksen suuntautuminen	Teknillisluonteisen koulutuksen indikaattori klt	0 tai 1
Tehtävä	79 tehtäväindikaattoria	0 tai 1
Yritys	125 yritysindikaattoria	0 tai 1
Sopimusala	37 sopimusalaindikaattoria	0 tai 1

3.3 Joitakin tilastollisia näkökohtia

Koska aineisto muodostaa kokonaisotannan, tuntuisi luontevalta sivuuttaa tavanomaiset tilastollisen merkitsevyyden testit ja pitää pienintäkin palkanmuodostuksessa havaittua piirrettä merkityksellisenä. Tähän argumenttiin ja tilastolliseen testaamiseen otettava kanta riippuu täysin siitä, mitä dataa generoivalla prosessilla (DGP) tarkoitetaan. Jos se ymmärretään suppeasti, pidetään DGP:na perusjoukosta eli teollisuussektorilla vuoden 2000 lopulla työskennelleistä toimihenkilöistä tehtyä otantaa. Ainoa DGP:iin liittyvä epävarmuustekijä on, että otannassa olisi voitu saada toisenlainen aineisto. Päättely koskee tällöin ainoastaan niitä henkilöitä, jotka sattuiivat olemaan kyseisessä perusjoukossa aineiston keruuhetkellä. Kun heistä on tehty kokonaisotanta, ei päättelyyn luonnollisestikaan liity mitään tilastollista epävarmuutta ja esitetty argumentti on oikeutettu. Jos DGP sen sijaan ymmärretään laajasti, pidetään kiinnostuksen kohteena koko sitä prosessia, joka tuottaa erilaiset palkat teollisuuden toimihenkilöille (Spanos 1986, 3–22). Otannan lisäksi epävarmuutta aiheuttaa toinenkin piirre: täysin sama DGP olisi voinut tuottaa erilaisen perusjoukon. Ajatellaan, että on tavallaan

tehty kaksi otantaa: ensin DGP synnyttää tietyllä ajanhetkellä jonkin perusjoukon, josta edelleen osa valitaan aineistoksi. Kokonaisotanta poistaa vain toisen epävarmuuden lähteen. Sen kautta syntynyttä aineistoa pidetään suurena otoksena superpopulaatiosta, ja aineisto on vain yksi sen monista mahdollisista ilmentymistä. Kun DGP ymmärretään näin, on perusteltua todentaa miesten ja naisten palkkayhtälöiden väliset erot tilastollisin testein.

Regressioajot tehtiin SAS-ohjelmiston versiolla 8,2. Koska White –tyyppisesti tehdyt testit hylkäsivät selvästi nollahypoteesin homoskedastisuudesta, korjattiin p-arvot heteroskedastisuuden suhteen Whiten esittämällä tavalla. Menetelmä on johdettu asymptoottisen teorian avulla, joten se toimii vain tarpeeksi suurissa aineistoissa. (White 1980, 817–838.) Sen soveltuvuutta arvioitaessa katsottiin, että tutkielman aineisto on riittävän iso ja täyttää tämän ehdon. Korjaus on tehty kaikille tutkielmassa raportoitaville p-arvoille, joten sitä ei mainita enää uudelleen estimointitulosten yhteydessä. Sen vaikutus oli yllättävän vähäinen: esimerkiksi kohdassa 4.3 esiteltävässä mallin supistuksessa olisi päädytty tavanomaisiakin p-arvoja käyttämällä täysin samoihin johtopäätöksiin. Havaintoa voidaan ainakin jossain määrin pitää osoituksena siitä, että heteroskedastisuus ei ole kovin voimakasta.

Kun tuonnempana arvioidaan muuttujien tilastollista merkitsevyyttä, pidetään sen rajana yhden prosentin p-arvoa tavanomaisen viiden prosentin sijasta. Syynä menettelyyn ei ole niinkään se, että haluttaisiin tehdä varovaisempaa päättelyä, vaan pikemminkin ratkaisuun päädyttiin aineiston laajuuden vuoksi. Suuressa aineistossa melkein mikä tahansa muuttuja on tilastollisesti merkitsevä palkan selittäjä, vaikka sen asiallinen merkitys tutkittavan ilmiön kannalta olisi olematon. Erityisen hyvin tämä tulee esiin kohdassa 4.3 esiteltävässä mallin supistuksessa. Vaikka mallista poistettiin monta viiden prosentin riskitasolla merkitsevää muuttujaa, ei sen kyky selittää toimihenkilöiden palkkaa tai sen antama arvio palkkadiskriminaatiosta muuttunut miksikään. Suuressa aineistossa jonkin muuttujan tilastollinen ja asiallinen merkitsevyys näyttäisivät olevan kaksi eri asiaa, joskin useimmiten ne yhtyvät. Ensin mainittu on jälkimmäisen välttämätön mutta ei riittävä ehto. Asian tarkempi pohtiminen tarkoittaisi oikeastaan uusien merkitsevyyden kriteerien kehittämistä, eikä se kuulu tämän tutkielman piiriin. Tilastollisessa päättelyssä pitäydyttiin tavanomaisessa ja tutkimusongelman kannalta ilmeisesti tarpeettoman varovaisessa menettelyssä. On kuitenkin pohtimisen arvoinen seikka, voisiko palkkadiskriminaation luotettavaan arviointiin käyttää yksinkertaisempaa mallia kuin tuonnempana esiteltävä PM1. Tutkielman tavoitehan ei suinkaan ole toimihenkilöiden palkanmuodostuksen mahdollisimman tarkka mallintaminen, vaan se on vain palkkasyrjinnän tutkimisen väline. Siinä tehdyt yksinkertaistukset ovat hyväksyttäviä, jos ne eivät muuta itse diskriminaatiosta tehtäviä johtopäätöksiä. Tämä pätee riippumatta siitä, ovatko yksinkertaistukset puhtaasti tilastollisin kriteerein arvioituina oikeutettuja vai eivät. Pohdintaa sovelletaan käytäntöön kohdassa 6.1, jossa esitellään PM1:tä radikaalisti yksinkertaistamalla konstruoitu palkkamalli GPM1. Huomionarvoista on,

että se antaa ilmiselvistä tilastollisesta virheellisyydestään huolimatta samat tulokset kuin sitä selvästi monimutkaisempi PM1.

Yleensä laajojen regressiomallien yksittäisiä kertoimia on hankala tulkita, koska selittävät muuttujat ovat voimakkaasti korreloituneita. Usein esitettävälle ceteris paribus –oletukselle ei välttämättä löydy konkreettista sisältöä. (Vrt. Koev 1996, 43.) Koska tutkielman estimoinneissa käytetään pienimmän neliösumman (PNS) menetelmää, voidaan kertoimien tulkinnassa käyttää apuna Frisch-Waugh-Lovell –lausetta (FWL-lause). Muodostetaan matriisi \mathbf{X} siten, että kaikki selittävät muuttujat ovat sen sarakkeina. Tällöin matriisi $\mathbf{P}_X \equiv \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ projisoi selitettävän \mathbf{y} ortogonaalisti matriisin \mathbf{X} sarakeavaruuteen ja yhtälöstä $\mathbf{P}_X\mathbf{y} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ saadaan kyseisen PNS-regression sovitteet. Residuaalit saadaan yhtälöstä $\mathbf{M}_X\mathbf{y} \equiv (\mathbf{I} - \mathbf{P}_X)\mathbf{y}$, jossa matriisi \mathbf{M}_X projisoi selitettävän \mathbf{y} matriisin \mathbf{X} sarakeavaruuden ortogonaaliin komplementtiin. Jaetaan selittäjäjoukko edelleen kahteen osaan, joiden muodostamia matriiseja merkitään \mathbf{X}_1 ja \mathbf{X}_2 . Niihin liittyvät projektiomatriisit $\mathbf{P}_1 \equiv \mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'$ ja $\mathbf{P}_2 \equiv \mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^{-1}\mathbf{X}_2'$ ja niitä vastaavat residuaalimatriisit \mathbf{M}_1 ja \mathbf{M}_2 muodostetaan kuten edellä. Käytetään ensin selittäjinä kaikkia muuttujia eli lasketaan regressio

$$(14) \quad \begin{aligned} \mathbf{y} &= \mathbf{X}\mathbf{b} + \text{residuaalit} \\ &= \mathbf{X}_1\mathbf{b}_1 + \mathbf{X}_2\mathbf{b}_2 + \text{residuaalit}. \end{aligned}$$

Tehdään seuraavaksi kaksi apuregressiota: selitetään muuttujajoukolla \mathbf{X}_1 kaikkia joukon \mathbf{X}_2 muuttujia, jolloin saadaan residuaaleina $\mathbf{M}_1\mathbf{X}_2$ \mathbf{X}_1 -puhdistetut \mathbf{X}_2 -muuttujat. Poistetaan myös projektiolla $\mathbf{M}_1\mathbf{y}$ riippuvasta muuttujasta \mathbf{y} muuttujilla \mathbf{X}_1 selittyvä osuus. FWL-lauseen sisältö on seuraava: regressiosta

$$(15) \quad \mathbf{M}_1\mathbf{y} = \mathbf{M}_1\mathbf{X}_2\mathbf{b}_2 + \text{residuaalit}$$

saadaan täsmälleen sama kerroinvektori \mathbf{b}_2 muuttujajoukolle \mathbf{X}_2 kuin regressiosta (14). Myös residuaalit ovat yhteneviä, ja ne vastaavat molemmissa tapauksissa numeerisesti jäännöksiä $\mathbf{M}_X\mathbf{y}$. Projektiot (14) ja (15) ovat siis identtisiä tapoja mallintaa sama ilmiö. (Davidson – McKinnon 1993, 19–23.)

FWL-lause johdettiin alun perin aikasarja-aineistoissa esiintyvän kausivaihtelun puhdistamista varten. Sen tulokset pätevät kuitenkin missä tahansa PNS-regressiossa (mts. 19–23), ja niinpä sitä sovelletaan tässä palkkamallin kertoimien tulkintaan. Jako kahteen selittäjäjoukkoon voidaan tehdä mielivaltaisesti. Regression (14) tyyppisessä laajassa palkkamallissa esiintyvän yksittäisen muuttujan ja sen saaman kertoimen merkitystä pohdittaessa on hyödyllisin erotelu siihen ja kaikkiin muihin selittäjiin. Yhtälön (15) mukaan tulkittuna kerroin kuvaa kyseisen muuttujan muihin selittäjiin nähden ortogonaalisen osan kykyä selittää palkan kaikkiin muihin muuttujiin nähden ortogonaalista osaa. Ainoas-

taan korreloitumattomien selittäjien tapauksessa voidaan kerroinvektorille \mathbf{b}_2 antaa itsenäinen muuttujajoukosta \mathbf{X}_1 riippumaton tulkinta. Tällöin pätee $\mathbf{M}_1\mathbf{X}_2 = \mathbf{X}_2$ ja regressiosta (14) saatava \mathbf{b}_2 -estimatti pysyy muuttumattomana, vaikka selittäjistä poistettaisiin muuttujajoukko \mathbf{X}_1 . Muulloin \mathbf{b}_2 :n merkitys on arvioitava yhtälön (15) mukaan.

4 Palkkamallin estimointi

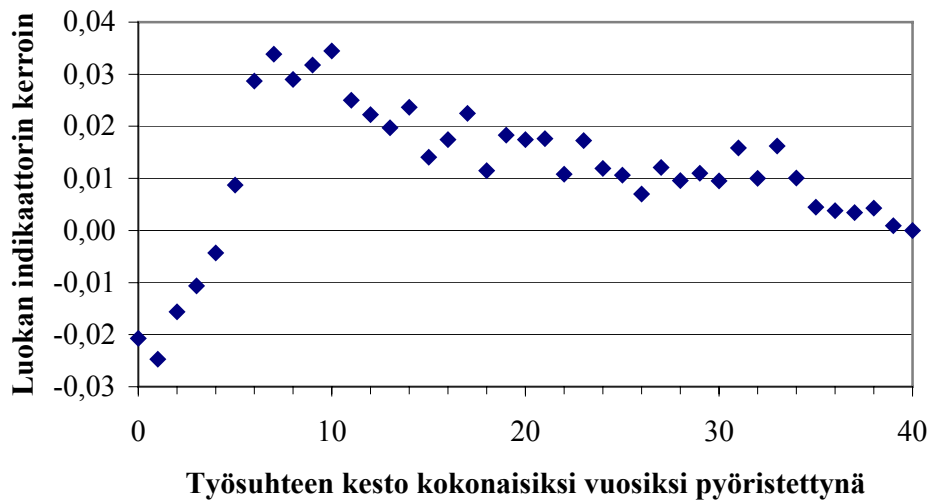
4.1 Jatkuvien muuttujien muunnosten valinta

Ennen toimihenkilöille muodostettavan palkkamallin estimointia oli päätettävä, mitä jatkuvien muuttujien muunnoksia otetaan selittäjiksi. Inhimillisen pääoman teorian (esim. Asplund 1993a) ja aiemman tutkimuksen (esim. Vartiainen 2001) perusteella on nimittäin aihetta odottaa, että mm. iän rajatuotto on laskeva. Piirre voidaan mallintaa useilla vaihtoehtoisilla funktiomuodoilla, ja tavanomainen ratkaisu on käyttää iän lisäksi selittäjänä sen neliötä (esim. Lilja 1997). Myös iän neliöjuuri tulee kyseeseen (esim. Koev 1996). Koska käsillä oleva aineisto on korkeatasoinen ja erittäin laaja, voitiin sopivin muunnos päätellä suoraan sen perusteella. Kuvailtaan seuraavaksi, kuinka tämä tehtiin.

Taulukossa 5 luetellut jatkuvat muuttujat yleisvaativuus, työsuhteen kesto, ikä ja koulutusvuodet muutettiin ensin diskreeteiksi. Ikä on TT:n aineistossa ilmoitettu kokonaisina vuosina, ja koulutukseen käytetyt vuodet voivat saada vain taulukossa 1 luetellut 6 arvoa, joten ne olivat jo valmiiksi luokitellussa muodossa. Työsuhteen kesto pyöristettiin yhden ja yleisvaativuus puolen yksikön tarkkuudella, jotta saatiin rajattua niiden arvot tiettyihin lukuihin. Kun kaikki jatkuvat muuttujat oli luokiteltu tällä tavalla, muodostettiin jokaiselle luokalle indikaattori. Tämä mahdollisti iän, työsuhteen keston, yleisvaativuuden ja koulutusvuosien käyttämisen diskreetteinä selittäjinä.

Koska selittäjät ovat keskenään voimakkaasti korreloituneita, ei mahdollisia epälineaarisia riippuvuuksia kannattanut tutkia selittämällä palkkaa vuorotellen yksittäisillä diskreeteiksi muunnetuilla muuttujilla. Toisen tai useamman muuttujan lisääminen malliin olisi muuttanut kertoimia FWL-lauseen havainnollistamalla tavalla. Siksi estimoitiin testimalli TPM, jossa huomioitiin keskeiset palkanmuodostuksen piirteet. Palkan logaritmia selitettiin kaikilla taulukon 5 muuttujilla, joista neljä jatkuvaa oli muutettu edellä kuvatulla tavalla diskreeteiksi. Niille sallittiin siis täysin vapaat palkkaprofiilit, joissa oli otettu kaikkien muiden tekijöiden vaikutus huomioon. Profiileja tarkasteltiin graafisesti merkitsemällä koordinaatiston vaaka-akselille suuruusjärjestykseen jatkuvan muuttujan luokat ja pystyakselille luokkien indikaattorien saamat kertoimet. Kuviossa 5 on esimerkkinä näin saatu työsuhteen keston palkkaprofiili.

Kuvio 5. Diskreetin työsuhteen keston palkkaprofiili mallissa TPM



Koska funktiomuotojen analyysi ei kuulu tutkielman ydinalueisiin, raportoidaan tässä vain keskeiset mallilla TPM saadut tulokset. Tärkein graafisten tarkastelujen kautta saatu tieto on, että yleisvaativuuden vaikutus palkan logaritmiin *säilyy lineaarisena*. Tämä ei ollut etukäteen mitenkään selvää, sillä ylva johdettiin työehtosopimuksien eikä empiiristen havaintojen perusteella. Riippuvuus olisi voinut laajempaa selittäjoukkoa käytettäessä muuttua epälineaariseksi tai kadota kokonaan muuttujien keskinäisen korreloituneisuuden vuoksi. Koulutusvuosien vaikutus palkan logaritmiin on myös – jopa hiukan yllättäen – lähes lineaarinen, joten kv8:aa käytetään selittäjänä sellaisenaan (vrt. Asplund 1993b, 66–73).

Työsuhteen keston vaikutus palkan logaritmiin on jo mutkikkaampi, kuten kuvio 5 ilmenee. Aineistoon parhaiten sopivan muunnoksen etsimisessä edettiin siten, että tysk vaihdettiin takaisin jatkuvaksi muuttujaksi ja lisättiin malliin sen neliö epälineaarisen riippuvuuden huomioimiseksi. Kaikki muut selittäjät pidettiin luokittelevina, joten malli erosi alkuperäisestä TPM:stä ainoastaan työsuhteen keston mallintamisen osalta. Kun se estimoitiin, saatiin työsuhteen kestolle jatkuva palkkaprofiili, jota verrattiin kuviossa 5 esitettyyn vapaaseen profiiliin. Näin saatiin kuva neliömuunnoksen soveltuvuudesta työsuhteen keston epälineaarisen palkkavaikutuksen mallintamiseen. Graafisen tarkastelun lisäksi toisena teknisen sopivuuden kriteerinä käytettiin selitysasteen muutosta alkuperäiseen TPM:ään verrattuna. Samalla tavalla testattiin muuttujan tysk neliöjuurimuunnos, ja havaittiin, että se ottaa neliötä paremmin huomioon kuvion 5 piirteet. Niinpä sitä päätettiin käyttää jatkossa selittäjänä.

Ennako-odotusten mukaisesti myös ikä vaikuttaa palkan logaritmiin epälineaarisesti: sen rajatuotto on selvästi laskeva. Sille tehtiin täsmälleen sama tarkastelu kuin työsuhteen kestolle ja neliöjuurimuunnos osoittautui senkin osalta neliötä

paremmaksi selittäjäksi. Samalla havaittiin, että 18 vuoden vähentäminen iästä parantaa entisestään palkkamallin sopivuutta aineistoon. Ilmiö johtuu siitä, että neliöjuurifunktio kaareutuu voimakkaasti heti nollan jälkeen, ja sama piirre esiintyy iän palkkaprofiilissa – tosin vasta aineiston alimman iän eli 18 vuoden jälkeen. Kun siirrytään käyttämään muuttujaa ikä18 ja sen neliöjuurimuunnosta, saadaan funktion alun kaareutuminen paremmin hyödynnettyä. Tämän ja kohdassa 3.2 mainitun vakion tulkittavuuden lisäksi toimenpiteellä saavutetaan kolmas-kin etu. Iän neliöjuurifunktio on vaihteluvälillä 18–65 melkein lineaarinen ja siksi lähes täydellisesti korreloitunut itse iän kanssa. Niiden välinen korrelaatiokerroin on 0,99800. Muuttujan ikä18 ja sen neliöjuuren välinen korrelaatiokerroin on 0,99029 eli hiukan alempi, koska neliöjuurifunktion epälineaarisin osa otetaan käyttöön. Skaalauksella saavutettava multikollineaarisuuden väheneminen on kylläkin tämän tutkielman kannalta lähestulkoon merkityksetön, koska aineiston laajuus mahdollistaa voimakkaasti korreloituneidenkin muuttujien käytön. Sen sijaan pienessä aineistossa se saattaa ratkaista, estimoituuko malli ollenkaan.

4.2 Laaja palkkamalli LPM

Luvussa 2 esiteltyä mallinnusfilosofiaa noudattaen toimihenkilöille muodostettiin yhteinen palkkamalli, jossa interaktioiden käytöllä sallittiin kertoimien vapaa vaihtelu sukupuolen ja toimihenkilöryhmän mukaan. Jälkimmäinen palkanmuodostuksen piirre mahdollistettiin, koska aiemmassa tutkimuksessa on viitteitä sen tarpeellisuudesta (Asplund 1993b, Vartia ym. 1992 ja Koev 1996). Estimoinnin lähtökohdaksi muodostettiin laaja palkkamalli LPM. Siinä palkan logaritmin selittäjinä olivat kaikki taulukossa 5 luetellut muuttujat ja lisäksi joukko niiden avulla muodostettuja interaktioita, jotka luetellaan taulukossa 6. Sen esitystapa on tilan säästämiseksi varsin lyhyt; kaikki interaktiot saadaan kertomalla sulut auki. Yhteensä niitä on 42.

Taulukko 6. Laajan palkkamallin LPM interaktiot

Ensimmäisen asteen interaktiot

(kot, tet, ylt)*(sqrt(ikä18), sqrt(tysk), kv8)

(kot, tet)*ylva

klt*kv8

n*(sqrt(ikä18), sqrt(tysk), kv8, ylva, ylt, tet, kot, ylvapt, kalleulk)

n*(10 suurinta sopimusala)

Toisen asteen interaktiot

n*(kot, tet, ylt)*(sqrt(ikä18), sqrt(tysk), kv8)

n*(kot, tet)*ylva

Monet sopimusalat ovat hyvin pieniä, joten miesten ja naisten erilaiset palkkatasot sallittiin kymmenellä suurimmalla alalla, jotka kattavat yhteensä 77,5 % aineistosta (ks. liite 1). Jatkuvien muuttujien muunnoksia valittaessa havaittiin, että iän ja työsuhteen keston palkkavaikutukset välittyvät pääasiassa niiden neliöjuurten kautta. Muuttujat itse takaavat mallissa lähinnä sen, että palkkaprofiilit kaareutuvat loppuosaltaan riittävän alas. Siksi katsottiin, että neliöjuurimuunnosten kanssa muodostetut interaktiot riittävät mallintamaan poikkeavat palkkaprofiilit aineiston eri ositteissa. Tehtävä- tai yritysindikaattorien kanssa ei muodostettu interaktioita niiden suuren määrän vuoksi, vaan niitä pidettiin lähinnä erilaisia palkkatasoja kontrolloivina apumuuttujina. Luvussa 6 tarkastellaan kyseisten interaktioiden merkitystä lähemmin. Palkkamallissa LPM on vakio mukaanluettuna 297 selittäjää.

4.3 Mallin LPM supistaminen malliksi PM1

Kun LPM estimoitiin, havaittiin, että monet sen interaktiot olivat tilastollisesti merkityksettömiä. Se sallii siis sellaisiakin palkanmuodostuksen piirteitä, joita ei oikeasti ole olemassa, mutta vasta sen estimoinnin jälkeen voitiin arvioida, miltä osin sitä oli mahdollista yksinkertaistaa. FWL-lauseen havainnollistamasta syystä kaikkia tilastollisesti merkityksettömiä muuttujia ei poistettu kerralla. Jo yhdenkin selittäjän pudottaminen mallista olisi saattanut muuttaa jonkin jäljelle jääneen alun perin merkityksettömän kertoimen tilastollisesti merkitseväksi. Siksi LPM supistettiin muuttuja kerrallaan. Siitä poistettiin ensin yksi tilastollisesti ei-merkitsevä selittäjä, ja malli estimoitiin uudelleen käyttäen jäljelle jääneitä muuttujia selittäjinä. Jos jokin niistä oli edelleen merkityksetön, se poistettiin, näin saatu uusi malli estimoitiin, ja menettely aloitettiin alusta. Tällä tavalla edettiin, kunnes kaikki jäljelle jääneet selittäjät olivat tilastollisesti merkitseviä. Karshintaa ei ulotettu tehtävä- yritys- ja sopimusalaindikaattoreihin niiden suuren määrän vuoksi.

Menetelmä ei välttämättä johda yksiselitteiseen lopputulokseen. Supistetun palkkamallin selittäjäjoukko saattaa riippua järjestyksestä, jossa muuttujia poistetaan. Siksi olisi pitänyt kokeilla useita vaihtoehtoisia reittejä yksinkertaistaa LPM. Koska laskenta-ajat olivat kuitenkin varsin pitkiä aineiston ja mallin laajuuden vuoksi, tehtiin supistus vain kerran. Niinpä poistettavia selittäjiä valittaessa käytettiin erityistä varovaisuutta. Supistaminen aloitettiin muuttujista, jotka kaikkein selkeimmin olivat tilastollisesti merkityksettömiä. Sitä jatkettiin mahdollisimman pientä aineiston osaa koskevilla jatkuvien muuttujien interaktioilla, koska mahdolliset niiden kohdalla tehdyt virhespesifikaatiot olisivat koskeneet pientä joukkoa. Erilaisia palkkatasoeroja kontrolloivat interaktiot pidettiin mahdollisesta tilastollisesta merkityksettömyydestään huolimatta mahdollisimman pitkään mukana. Jos jonkin jatkuvan muuttujan poisto olisi ollut liiallinen yksinkertaistus, olisivat ne korjanneet tilannetta. Samasta syystä naisindikaattori poistettiin vasta

viimeisenä, vaikka se oli tilastollisesti täysin merkityksetön koko supistamisen ajan.

Täysin ratkaisevaa jonkin selittäjän merkitystä arvioitaessa on, muuttaako sen pudottaminen mallista johtopäätöksiä. Siksi jokaisen poiston jälkeen laskettiin kyseisen yksinkertaistuksen seurauksena saadun mallin antama arvio palkkadiskriminaatiosta. Se laskettiin toimihenkilöiden keskimääräisen työtilanteen kohdalla. Laskutekniikka esitellään tarkemmin kohdassa 5.1, mutta arviota käytettiin jo tässä vaiheessa eräänlaisena mallin supistuksen kuumemittarina. Jos se olisi muuttunut selvästi, oltaisiin voitu päätellä, että mallista ollaan poistamassa tutkimustavoitteen kannalta olennaista informaatiota sisältävää muuttujaa. Myös selitystason muutosta seurattiin.

Koska mallin supistamisessa käytettiin subjektiivista harkintaa, esitetään taulukossa 7 sen vaiheet kokonaisuudessaan lukijan arvioitavaksi. Vasemmanpuoleisin sarake kertoo muuttujien poistamisjärjestyksen. Sarake $P(\text{ChiSq})$ kertoo muuttujan (White-tyyppisesti korjatun) p-arvon mallissa, jossa se viimeiseksi oli mukana. $PDF(\bar{\mathbf{x}})$ on palkkadiskriminaatio tyypillisen toimihenkilön kohdalla, kun sen arviointiin käytettiin kyseistä mallia. Viimeinen sarakkeessa on sen selitystaste. Esimerkiksi taulukon ensimmäisen rivin kaksi viimeistä lukua koskevat laajaa, supistamatonta palkkamallia. Siitä poistettiin ensimmäisenä muuttuja $n \cdot \text{tet} \cdot \sqrt{\text{tysk}}$, jonka p-arvo LPM:ssä oli 0,6115. Näin saatiin malli, jota koskevat toisen rivin tiedot, ja josta vuorostaan poistettiin muuttuja $n \cdot \text{kot} \cdot \sqrt{\text{tysk}}$. Viimeinen rivi kertoo supistamisen lopputuloksena saadun mallin PM1 selitystasteen ja sen antaman arvion palkkadiskriminaatiosta tyypillisessä toimihenkilötehtävässä.

Taulukko 7. Mallin LPM supistaminen malliksi PM1

Poistettu muuttuja	$P(\text{ChiSq})$	$PDF(\bar{\mathbf{x}})$	$R^2(n\text{-adj.})$
$n \cdot \text{tet} \cdot \sqrt{\text{tysk}}$	0,6115	0,0850	0,778436
$n \cdot \text{kot} \cdot \sqrt{\text{tysk}}$	0,4906	0,0850	0,7784
$n \cdot \text{kot} \cdot \text{kv}8$	0,7896	0,0849	0,7784
$n \cdot \text{kot} \cdot \sqrt{\text{ikä}18}$	0,1673	0,0849	0,7784
$n \cdot \text{ylt} \cdot \sqrt{\text{tysk}}$	0,2059	0,0850	0,7784
$\text{tet} \cdot \sqrt{\text{tysk}}$	0,8868	0,0850	0,7784
$\text{tet} \cdot \sqrt{\text{ikä}18}$	0,1629	0,0850	0,7784
$\text{kot} \cdot \text{kv}8$	0,0554	0,0850	0,7784
$\text{kot} \cdot \sqrt{\text{tysk}}$	0,9524	0,0850	0,7784
$n \cdot \text{tet}$	0,0474	0,0850	0,7784
$n \cdot \text{kalleulk}$	0,0124	0,0852	0,7784
$\text{ylt} \cdot \text{kv}8$	0,0301	0,0851	0,7784
$n \cdot \sqrt{\text{tysk}}$	0,0105	0,0851	0,7784
ylt	0,1229	0,0853	0,7784
n	0,9635	0,0853	0,7784
-	-	0,0853	0,778380

Siirtyminen mallista LPM malliin PM1 laski selitysasastetta 5,6 prosenttiyksikön tuhannesosaa. Muutos oli niin vähäinen, ettei se ilmennyt lainkaan regressio-ohjelman tavanomaisesta neljä desimaalia sisältävästä tulosteesta. Taulukossa raportoidut mallien LPM ja PM1 selitysasasteet jouduttiin tulostamaan kuuden desimaalin tarkkuudella erilliseen tiedostoon, jotta niitä voitiin vertailla. Näin kävi, vaikka käytettiin tavanomaista tiukempaa yhden prosentin riskitasoa. Voidaan päätellä, että data on erittäin redundanttii: poistetuilla muuttujilla ei ole juuri ollenkaan muista muuttujista riippumatonta kykyä selittää toimihenkilöiden palkkaa. Mallin LPM supistaminen malliksi PM1 ei myöskään muuttanut mitenkään arviota palkkadiskriminaatiosta, ja PM1:tä päätettiinkin käyttää jatkossa palkkasyrjinnän arviointiin. Taulukossa 8 esitellään mallin kertoimet ja niiden p-arvot. Sopimusala- yritys- ja tehtäväindikaattoreille estimoituja kertoimia ei raportoida niiden lukuisuuden vuoksi.

Taulukko 8. Palkkamalli PMI

<i>Muuttuja</i>	<i>Kerroin</i>	<i>t-arvo</i>	<i>P(ChiSq)</i>
vakio	1,89307	168,59	<0,0001
ikä18	-0,01174	-34,85	<0,0001
sqrt(ikä18)	0,14658	48,14	<0,0001
ylt*sqrt(ikä18)	0,06468	89,54	<0,0001
kot*sqrt(ikä18)	0,01225	9,28	<0,0001
n*sqrt(ikä18)	-0,02004	-25,95	<0,0001
n*tet*sqrt(ikä18)	0,00574	3,41	<0,0001
n*ylt*sqrt(ikä18)	-0,01257	-7,52	<0,0001
tysk	-0,00359	-17,14	<0,0001
sqrt(tysk)	0,03275	22,78	<0,0001
ylt*sqrt(tysk)	-0,00366	-4,73	0,0001
kv8	0,02027	47,10	<0,0001
klt	-0,09476	-39,97	<0,0001
klt*kv8	0,01840	37,70	<0,0001
tet*kv8	-0,00898	-9,50	<0,0001
n*kv8	-0,00579	-9,56	<0,0001
n*tet*kv8	0,00665	3,66	<0,0001
n*ylt*kv8	0,01269	14,84	<0,0001
ylva	0,04931	89,30	<0,0001
tet*ylva	-0,01020	-17,66	<0,0001
kot*ylva	-0,00609	-5,99	<0,0001
n*ylva	0,01216	15,46	<0,0001
n*kot*ylva	-0,01729	-13,93	<0,0001
n*tet*ylva	-0,00893	-5,74	<0,0001
ylvapt	0,28398	82,66	<0,0001
n*ylvapt	-0,06278	-15,31	<0,0001
kalleulk	0,03685	34,70	<0,0001
tet	0,14863	27,95	<0,0001
kot	-0,02619	-3,32	0,0051
n*kot	0,07500	11,98	<0,0001
n*ylt	0,05640	6,19	<0,0001
37 sopimusalaindikaattoria			
n*posti	0,03613	9,62	<0,0001
n*teleliikenneala	-0,03732	-7,70	<0,0001
n*tietoliikenneala	-0,02329	-5,00	<0,0001
n*suunn.- ja konsulttiala	0,02520	0,01	<0,0001
n*graafinen teoll.	-0,03857	-7,71	<0,0001
n*met. ja elektr. teoll.	0,01363	5,19	<0,0001
n*paperiteoll.	-0,05972	-13,53	<0,0001
n*talonrakennusala	0,02372	3,67	<0,0001
n*muovi- ja kemian tuotteet.	0,01909	4,24	<0,0001
n*meijerit	0,08012	15,58	<0,0001
125 yritysindikaattoria			
79 tehtäväindikaattoria			

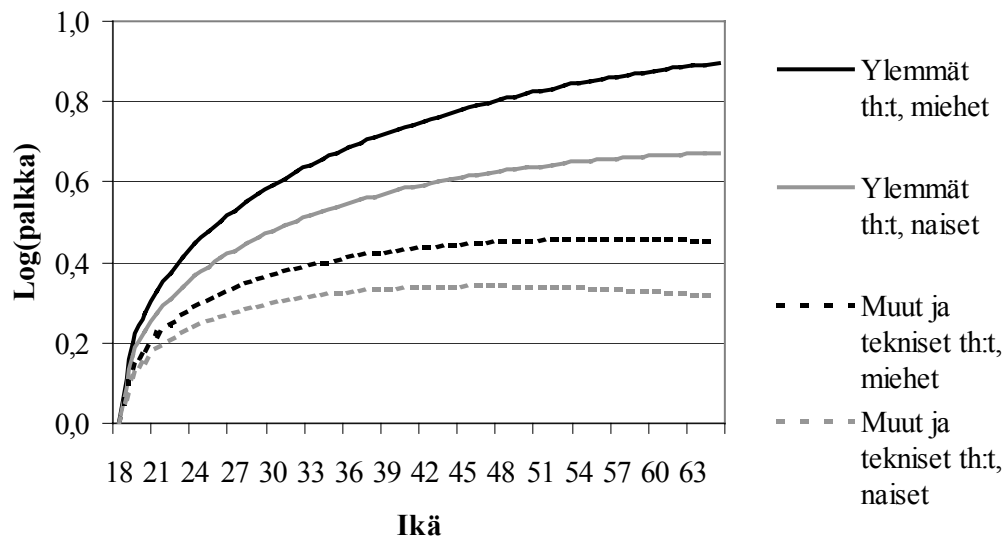
4.4 Toimihenkilöiden palkanmuodostuksen arviointia

Tarkastellaan seuraavaksi lähemmin muutamaa mallin PM1 keskeistä piirrettä ja havainnollistetaan sukupuolten erilaista kohtelua jatkuvien muuttujien osalta graafisesti. On syytä huomioida, että esiteltäville palkkaprofiileille voidaan antaa ainoastaan FWL-lauseen mukainen tulkinta. Kyseessä on aina toimihenkilön ominaisuuden tai työn piirteen residuaalivaikutus, josta on puhdistettu muiden muuttujien selityskyky.

Muuttujien skaalauksen ansiosta mallin PM1 regressiovakiolle löytyy mielekäs tulkinta. Se kuvaa palkkausta sellaisten nuorten ja matalasti koulutettujen toimihenkilöiden ryhmässä, jotka ovat alhaisen vaativuustason tehtävissä ja joiden työsuhde on aivan alussa. Se on eräänlainen arvio teollisuuden toimihenkilöiden lähtöpalkasta. Kun vakioista otetaan eksponenttimuunnos, saadaan sen suuruudeksi 6,64 euroa. Naisindikaattori putosi vastoin ennako-odotuksia pois tilastollisesti merkityksettömänä, eikä sen kautta saatu näyttöä naisten alkupalkkaedusta tai -haitasta. Vaikuttaa siltä, että miehet ja naiset aloittavat uransa samasta tilanteesta ja että palkkadiskriminaatio kehittyy vasta uran edetessä. Kuva saattaa tosin muuttua, kun otetaan huomioon muutkin palkkasyrjintään vaikuttavat tekijät.

Iän palkkavaikutus riippuu mallissa PM1 varsin mutkikkaalla tavalla sukupuolesta ja toimihenkilöryhmästä. Hallitseva ja kaikissa toimihenkilöryhmissä toistuva piirre on naisten miehiä selvästi alempi iän neliöjuurimuunnoksen kerroin: muuttujan $n \cdot \sqrt{\text{ikä} - 18}$ kerroin on negatiivinen. Naisia palkitaan siis iän tuomasta kokemuksesta miehiä huonommin (vrt. Koev 1996, 41–42). Tämä ilmenee hyvin kuviosta 6, jossa esitetään neljä PM1:n perusteella laskettua palkkaprofiilia. Niiden muodot poikkeavat esimerkiksi Koevin estimoimista (1996, 42), mikä johtuu kohdassa 4.1 perustellusta neliöjuurifunktion kaartuvan osan hyödyntämisestä. Sukupuolesta ja toimihenkilöryhmästä riippumatta palkka kasvaa nopeasti heti työuran alussa, mutta iän ansioita nostava vaikutus tasaantuu nopeasti ja kääntyy alemmilla toimihenkilöillä jopa laskuun. Naisilla tämä tapahtuu aiemmin kuin miehillä, ja erilaisesta kohtelusta aiheutuva sukupuolten välinen palkkaero kasvaa jatkuvasti iän myötä. Työuran lopulla se on ylempien toimihenkilöiden keskuudessa noin 20 log-% ja teknisten ja muiden toimihenkilöiden muodostamassa ryhmässä hiukan yli 10 log-%. Seuraavassa luvussa esitetään, kuinka tämä palkanmuodostuksen piirre osoittautui selvästi suurimmaksi palkkadiskriminaation lähteeksi.

Kuvio 6. Iän palkkavaikutus



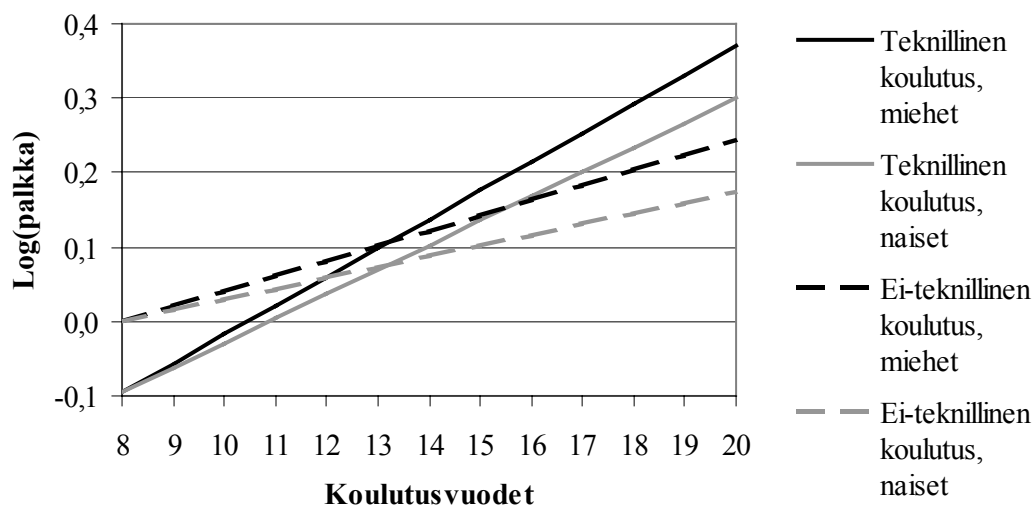
Kuvion 6 tulkinassa on otettava huomioon mallin PM1 laaja selittäjäjoukko. Tehtäväindikaattorien ja vaativuutta kuvaavan muuttujan käytön vuoksi iän palkkavaikutukselle jää FWL-lauseen havainnollistama residuaalinen rooli, josta on puhdistettu kyseiset työn piirteet. Koska uralla eteneminen ymmärretään yleensä juuri siirtymisenä paremmin palkattuihin ja haastavampiin tehtäviin, ei kuviossa 6 havainnollistettua miesten ja naisten erilaista ikäkohtelua voida niinkään pitää merkinä sukupuolten erilaisista urakehityksistä. Etenemisen vaikutus palkkaan tulee jo otetuksi huomioon edellä mainituissa muuttujissa – tässä vertaillaan samassa tehtävässä toimivia miehiä ja naisia. Ennemminkin tulee kyseeseen tulkinta, jonka mukaan sukupuolia palkitaan *samasta* etenemisestä eri tavalla: miehillä se kanavoituu syystä tai toisesta suuremmaksi palkan nousuksi kuin naisilla.

Ylennykset nostavat eniten teollisuuden toimihenkilöiden ansioita (Lilja 1996, 121), ja toisaalta miesten urakehitys on nopeampaa kuin naisten (Lilja 1997, 27). Tuntuu luontevalta olettaa, että miesten nopeampaan urakehitykseen liittyisivät myös useammat työnantajan vaihdot. Onkin yllättävää, ettei mallissa PM1 työsuhteen keston palkkavaikutus riipu sukupuolesta. Voidaan nimittäin arvioida, että osa työpaikan vaihdoksista tehdään sen kautta saatavan isomman palkan vuoksi. Jos miehillä olisi näitä vaihdoksia enemmän, liittyisi heillä lyhyeen työsuhteeseen korkeampi palkka kuin naisilla. Nyt tällaista piirrettä ei löydetty. Työsuhteessa pysyminen nostaa aluksi ansioita, mutta sitoutuminen samaan työnantajaan muuttuu lopulta rasitteeksi. Ei ole selvää, mistä ilmiö johtuu. Ilmeisesti toimihenkilön kyky hoitaa tehtävänsä paranee ensin oppimisen vuoksi, mutta jos toimeen juututaan liian pitkäksi aikaa, jäävät palkankorotukset saamatta. Mm. Lilja raportoi samansuuntaisia tuloksia tehtävän vaativuuden osalta (1996, 120–121). Haitta liian pitkästä työsuhteesta on sukupuolille yhtäläinen.

Koulutuksen palkkavaikutus riippuu niinkään monesta tekijästä. Jotta saataisiin esiin eräs kiinnostava piirre, esitetään kuviossa 7 palkkaprofiilit muiden ja konttoritoimihenkilöiden ryhmässä. Ennako-odotusten mukaisesti teknillisestä koulutuksesta palkitaan paremmin, tosin siihen liittyy matalampi lähtötaso. Tämä poikkeaa mm. Koevin (1996, 31) ja Vartian ja Kurjenojan (1992, 36–37) saamista tuloksista. Lisäksi naisia palkitaan huonommin kouluttautumisesta, mutta sukupuoli vaikuttaa palkkaprofiileihin yllättäen vähemmän kuin koulutuksen suuntautuminen. Siksi on tärkeää, että mallissa PM1 sallitaan muuttujan kv8 kertoimen vaihtelu koulutusalan mukaan. Ilman tätä ominaisuutta malli yliarvioisi palkkadiskriminaation kyseisen selittäjän osalta. Miesten kerroin estimoitaisiin todellista suuremmaksi, koska 70 % heistä on hankkinut teollisuudessa hyödyllisen teknillisen koulutuksen (taulukko 2). Sen parempaa tuottoa ei kuitenkaan voida pitää palkkadiskriminaationa, koska sukupuolilla on yhtäläiset mahdollisuudet hakeutua mieleiselleen koulutusalueelle.

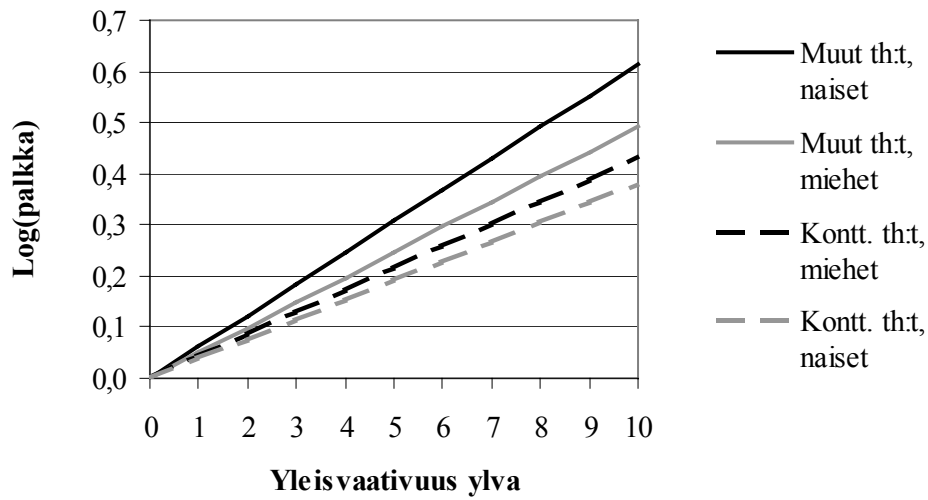
Yleisesti yksi lisävuosi koulutusta johtaa noin kahden log-prosentin suuruiseen palkanlisäykseen, mikä on aiempia tutkimuksia selvästi alempi arvio (esim. Asplund 1993b, 66–67). Ero johtuu tutkielman huomattavan laajasta selittäjäjoukosta: kahden prosentin nousussa on otettu muiden muuttujien vaikutus FWL-lauseeseen havainnollistamalla tavalla huomioon.

Kuvio 7. Koulutuksen vaikutus palkkaan muiden ja konttoritoimihenkilöiden keskuudessa



Työn vaativuus vaikuttaa voimakkaasti ansioihin. Toimihenkilön siirtyminen häntä koskevan asteikon alimmalta portaalta ylimmälle johtaa palkan kasvuun keskimäärin 40 log-prosentilla, mikä ilmenee kuvioista 8. Muiden toimihenkilöiden ryhmässä naisia palkitaan yllättäen vaativuustason noususta miehiä paremmin, ja vaikeimmissa tehtävissä ero nousee yli 10 log-prosentin. Konttoritoimihenkilöillä tilanne on päinvastainen, joskin ero on pienempi.

Kuvio 8. Yleisvaativuuden vaikutus palkkaan



Loput mallin PM1 muuttujat ovat luokittelevia, ja ne kontrolloivat aineiston eri ositteissa vaihtelevia palkkatasoja. Niiden kertoimille on hankala antaa graafista tulkintaa, eikä niitä pohdita tässä erikseen. Kyseisten muuttujien merkitys tulee esiin seuraavassa luvussa palkkadiskriminaation arvioinnin yhteydessä.

5 Palkkadiskriminaation tarkastelu

5.1 Empiirinen palkkadiskriminaatiofunktio

Mallin PM1 estimointituloksista saadaan miesten ja naisten palkkayhtälöt, joiden erotuksena kohdan 2.1 palkkadiskriminaatiofunktio (3) määritellään. Suurimmalla osalla muuttujista on mallissa sama kerroin miehille ja naisille, joten niiden kerroin PDF:ssa on nolla, ja ne supistuvat pois siitä. Siihen jäävät argumenteiksi vain ne muuttujat, joiden palkkavaikutus vaihtelee sukupuolen mukaan, ja niiden kertoimet PDF:ssa ovat muotoa $n \cdot xxx$ olevien interaktioiden vastalukuja. Esimerkiksi muuttujan $\sqrt{\text{ikä18}}$ kerroin miesten yhtälössä on 0,14658 ja naisten yhtälössä 0,12654. Jälkimmäinen saadaan lisäämällä ensimmäiseen interaktion $n \cdot \sqrt{\text{ikä18}}$ kerroin -0,02004. Muuttujan $\sqrt{\text{ikä18}}$ kerroin PDF:ssa on $0,14658 - 0,12654 = 0,02004$. Yhteensä mallissa PM1 on 22 muuttujaa, joiden kohdalla palkanmuodostus vaihtelee sukupuolen mukaan. Ne ovat PDF:n 22 argumenttia. Merkitään kyseisiä muuttujia vektorilla \mathbf{x} ja PDF:n kertoimia vektorilla \mathbf{b}'_{pdf} . Tällöin palkkadiskriminaatiofunktio on $PDF(\mathbf{x}) = \mathbf{b}'_{pdf} \mathbf{x}$ ja sen sama arvo riippuu siitä, mitä työtilannetta \mathbf{x} tarkastellaan.

Taulukossa 9 luetellaan mallista PM1 johdetut PDF:n argumentit ja niiden kertoimet. Samalla on laskettu esimerkinomaisesti jo kohdassa 4.3 hyödynnetty palkkadiskriminaatio toimihenkilöiden keskimääräisen työtilanteen kohdalla. Kolmas sarake kertoo kunkin muuttujan keskiarvon koko aineistossa. Neljäs sarake ilmaisee yksittäisen muuttujan absoluuttisen vaikutuksen palkkadiskriminaatioon. Se on toisen ja kolmannen sarakkeen alkioden tulo. Kun kaikkien 22 muuttujan syrjintävaikutukset lasketaan yhteen, saadaan alimman rivin arvio palkkadiskriminaatiosta tyypillisessä toimihenkilötehtävässä. Viimeiseen sarakkeeseen on laskettu kunkin argumentin suhteellinen osuus palkkasyrjinnästä, ja muuttujat on laitettu suuruusjärjestykseen sen itseisarvon mukaan.

Miesten ja naisten erilainen ikäkohtelu aiheuttaa selvästi suurimman osan palkkadiskriminaatiosta. Ei kuitenkaan voida sanoa, että syrjintä katoaisi tai muuttuisi peräti negatiiviseksi, jos sukupuolia palkittaisiin iästä yhtäläisesti. Muuttujilla ei ole itsenäistä muista selittäjistä riippumatonta tulkintaa, ja suhteellisten vaikutusten laskemisella tavoitellaan lähinnä kuvaa siitä, mitkä sukupuolten palkanmuodostusten väliset erot ovat keskeisiä syrjinnän lähteitä. Esimerkiksi alojen sisäiset palkkatasoerot vaikuttavat diskriminaatioon varsin vähän.

Taulukko 9. *Empiirinen palkkadiskriminaatiofunktio*

Muuttuja	Muuttujan ker- roin PDF:ssa	Muuttujan keski- arvo aineistossa	Absol. vaikutus palk- kadiskriminaatioon	Absol. vaikutuksen osuus pd:sta, %
sqrt(ikä18)	0,02004	4,68394	+0,0939	+110
ylvapt	0,06278	0,51997	+0,0326	+38
ylva	-0,01216	2,50969	-0,0305	-36
ylt*kv8	-0,01269	2,24864	-0,0285	-33
kv8	0,00579	4,53381	+0,0263	+31
ylt*sqrt(ikä18)	0,01257	1,68370	+0,0212	+25
ylt	-0,05640	0,37209	-0,0210	-25
kot	-0,07500	0,10397	-0,0078	-9
kot*ylva	0,01729	0,37092	+0,0064	+8
tet*ylva	0,00893	0,67417	+0,0060	+7
met. ja elektr. t.	-0,01363	0,34563	-0,0047	-6
posti	-0,03613	0,10398	-0,0038	-4
tet*sqrt(ikä18)	-0,00574	0,63158	-0,0036	-4
tet*kv8	-0,00665	0,54485	-0,0036	-4
paperiteoll.	0,05972	0,05585	+0,0033	+4
meijerit	-0,08012	0,03054	-0,0024	-3
teleliikenneala	0,03732	0,04294	+0,0016	+2
graafinen teoll.	0,03857	0,04011	+0,0015	+2
tietoliikenneala	0,02329	0,04198	+0,0010	+1
suunn. ja kons. ala	-0,02520	0,03779	-0,0010	-1
muovi- ja kemian tuotet.	-0,01909	0,04389	-0,0008	-1
talonrakennusala	-0,02372	0,03258	-0,0008	-1
Yhteensä			0,0853	101

Taulukossa 9 laskettiin palkkadiskriminaatio toimihenkilöiden tyypillisen työtilanteen kohdalla sijoittamalla PDF:n argumentiksi kyseistä työtilannetta kuvaava 22-ulotteinen vektori. Vastaavasti voitaisiin laskea palkkasyrjintä minkä tahansa aineiston ositteen keskimääräisessä työtilanteessa. Koska PDF on lineaarinen argumenttiensa suhteen⁶, päästään samaan tulokseen toistakin reittiä pitkin. Olkoon jossakin aineiston ositteessa n toimihenkilöä, ja kuvatkoon ositteen yksittäisen toimihenkilön työtilannetta vektori \mathbf{x}_i . Tällöin palkkadiskriminaatio ositteen tyypillisessä työtilanteessa $\bar{\mathbf{x}}$ on

$$(16) \quad PDF(\bar{\mathbf{x}}) = \mathbf{b}'_{pdf} \bar{\mathbf{x}} = \mathbf{b}'_{pdf} n^{-1} \sum_{i=1}^n \mathbf{x}_i = n^{-1} \sum_{i=1}^n \mathbf{b}'_{pdf} \mathbf{x}_i.$$

Palkkasyrjintä jonkin ositteen keskimääräisessä työtilanteessa saadaan siis myös laskemalla PDF:n avulla palkkadiskriminaatio jokaiselle toimihenkilölle juuri hänen työtilanteessaan ja laskemalla niiden keskiarvo ositteessa. (Vartia ym. 1992, 82–83.) Tämä on helpompi kuin taulukossa 9 esitelty tapa, ja lisäksi hen-

⁶ Vaikka iän ja työsuhteen keston palkkavaikutukset ovat epälineaarisia, ovat PM1 ja siitä johdettu PDF lineaarisia muuttujien sqrt(ikä18) ja sqrt(tysk) suhteen.

kilökohtaista palkkasyrjintäarviota voidaan tarkastella kuin mitä tahansa muutujaa. Analyysia voidaan viedä eteenpäin käyttämällä sitä vaikkapa regressiomallissa joko selittäjänä tai selitettävänä. Loput luvun tulokset laskettiin juuri tällä tavalla henkilötasolta lähtien. Niiden avulla saadaan kuva siitä, keitä toimihenkilöitä palkkadiskriminaatio erityisesti koskettaa ja siitä, miten syrjintä kehittyi työntekijän ominaisuuksien tai tehtävän piirteiden muuttuessa.

5.2 Palkkasyrjintä teollisuuden eri osissa

Taulukossa 10 esitellään diskriminaatioiden keskiarvot miesten ja naisten osajoukoissa. Ne vastaavat siis tarkalleen miesten ja naisten tyypillisissä työtilanteissa laskettuja syrjintäarvioita. Vastaavat luvut raportoidaan melkein aina muissa tutkimuksissa, joten ne muodostavat vertailukohdan. Lisäksi taulukkoon on laskettu palkkasyrjinnän pienin ja suurin arvo sekä sen ensimmäinen ja kolmas kvartiili, jotka kuvailevat diskriminaation jakaumaa kussakin aineiston ositteessa.

Taulukko 10. Palkkadiskriminaatio sukupuolen mukaan

<i>Osite</i>	<i>Min.</i>	<i>Q1</i>	<i>Keskiarvo</i>	<i>Q3</i>	<i>Max.</i>
Koko aineisto	-0,0899	0,0465	0,08526	0,1216	0,2963
Miehet	-0,0733	0,0453	0,08396	0,1203	0,2963
Naiset	-0,0899	0,0493	0,08746	0,1253	0,2705

Palkkadiskriminaatio toimihenkilöiden keskimääräisessä työtehtävässä on kahdeksan ja puoli log-prosenttia, mikä on melko tarkkaan kolmannes sukupuolten keskipalkkojen 25 log-prosentin suuruudesta erosta. Loppuosa eli kaksi kolmannesta palkkaerosta johtuu miesten ja naisten erilaisista henkilökohtaisista ominaisuuksista ja eroista heidän työnsä piirteissä. Palkkadiskriminaatioarviot ovat selvästi alempia kuin aiemmissa tutkimuksissa saadut teollisuuden toimihenkilöitä koskevat tulokset. Koev raportoi syrjinnän suuruudeksi 12,3 log-% miesten ja 10,6 log-% naisten tyypillisessä työtilanteessa (1996, 45), ja Vartian ja Kurjenojan tutkimuksen vastaavat luvut ovat 14,3 log-% ja 8,9 log-% (1992, 83). Jälkimmäisessä tosin tarkastellaan yli 200 hengen suuruisia metsä- ja metalliteollisuuden yrityksiä, joten tulokset eivät ole aivan suoraan vertailukelpoisia. Asplund raportoi palkkasyrjinnän suuruudeksi yli 20 prosenttia⁷ (1996, 3 ja 8). Ainakin osittain tutkielman diskriminaatioarvion pienuus palautuu luvun 1 pohdintaan. Sama työ on tässä määritelty lähes 300 muuttujan avulla, mikä on kriteerinä tiukimmasta päästä. Vastaavasti näkemys palkkasyrjinnästä on suppea, kuten analysoitava ansiokäsitekin. Toisaalta viitatus tutkimustulokset perustuvat vanhempiin aineistoihin, ja tilanne on hyvinkin voinut muuttua tasa-arvoisempaan

⁷ Kyseessä ovat tavalliset prosentit.

suuntaan. Laajemman selittäjäjoukon vaikutuksesta saadaan käsitys, kun verrataan tuloksia Korkeamäen ja Kyyrän 11 log-prosentin suuruiseen arvioon palkkasyrjinnästä. He analysoivat myös suppeaa ansiokäsitettä ja tutkimus perustuu samaan teollisuuden toimihenkilöiden palkka-aineistoon vuodelta 2000. Inhimillistä pääomaa kuvaavat muuttujat ovat tutkimuksessa jokseenkin samat kuin tässäkin, mutta siinä ei kontrolloida työn vaativuutta. (Korkeamäki ym. 2002, 1–15.) Näyttäisi siltä, että kun saman työn käsitteeseen sisällytetään yhtenevä työn vaativuus, saadaan alempia palkkasyrjintäarvioita.

Palkkadiskriminaatio vaihtelee melko paljon. Puolella toimihenkilöistä se sijoittuu välille 4,7–12,2 log-prosenttia. Suurimmillaan se on lähes 30 log-%, mutta joillakin miestoimihenkilöillä se on taas negatiivinen. He saavat matalampaa palkkaa kuin samoilla taustaominaisuuksilla varustetut naiset. Aineistosta löytyy myös naisia, joiden ansiot ylittävät vastaavaa työtä tekevien miesten palkat melkein yhdeksällä log-prosentilla. Syrjintä naisille tyypillisessä työtilanteessa on korkeampi kuin vastaava arvio miehille. Lähestulkoon jokaisessa edellä viitatussa tutkimuksessa suuruusjärjestys on päinvastainen ja lisäksi kahden arvion välinen ero on niissä johdonmukaisesti suurempi. Koska ne ovat hyvin lähellä toisiaan, raportoidaan jatkossa selkeyden vuoksi ainoastaan palkkadiskriminaatio kunkin ositteen keskimääräisessä työtilanteessa.

Taulukosta 11 havaitaan, että palkkasyrjintä on suurempaa ylempien kuin alemmien toimihenkilöiden keskuudessa. Koev raportoi samansuuntaisia tuloksia (1996, 47). Tässäkin on otettava huomioon, että jälkimmäisessä ryhmässä sama työ on aineistosta johtuen määritelty tiukemmin kuin ensimmäisessä: ylempien toimihenkilöiden osalta ei ole voitu kontrolloida työn vaativuutta. Niinpä osa ryhmän korkeammasta palkkadiskriminaatiosta saattaa johtua siitä, että tarkasteltava käsite on puuttuvan informaation vuoksi väistämättä väljempi. Teknisten toimihenkilöiden keskuudessa esiintyy vähiten palkkasyrjintää. Koev (mts. 47) ja Asplund (1996, 3) päätyvät samaan tulokseen, joskin he arvioivat itse diskriminaation suuremmaksi. Syrjintä on yllättävän suurta konttoritoimihenkilöillä. Koska kvartiilien erotus Q3 - Q1 on ryhmässä pienin, on palkkadiskriminaatio lisäksi voimakkaimmin keskittynyt ositteen painopisteen ympärille.

Taulukko 11. Palkkadiskriminaatio toimihenkilöryhmittäin

<i>Toimihenkilöryhmä</i>	<i>Min.</i>	<i>Q1</i>	<i>Keskiarvo</i>	<i>Q3</i>	<i>Max.</i>
Ylemmät th:t	-0,0367	0,0699	0,10497	0,1354	0,2780
Tekniset th:t	-0,0747	0,0266	0,06176	0,0934	0,2154
Konttorith:t	-0,0899	0,0698	0,09891	0,1273	0,2221
Muut th:t	-0,0606	0,0314	0,07073	0,1041	0,2963

Taulukossa 12 esitetään keskimääräiset palkkadiskriminaatiot kymmenellä suurimmalla sopimusallalla (ks. liite 1). Syrjintä on suurinta graafisessa ja paperi-

teollisuudessa, eikä jälkimmäisestä löydy ainuttakaan negatiivista arvoa. Alhaisinkin lukema on 6,4 log-%. Vähiten palkkadiskriminaatiota esiintyy talonrakennus- ja meijerialalla, ja niillä myös syrjinnän maksimi-arvot jäävät mataliksi. Havainnot ovat linjassa Koevin tulosten kanssa (mts. 48).

Taulukko 12. Palkkadiskriminaatio suurimmilla sopimusaloilla

<i>Sopimusala</i>	<i>Min.</i>	<i>Q1</i>	<i>Keskiarvo</i>	<i>Q3</i>	<i>Max.</i>
Met.- ja elektr. teoll.	-0,0606	0,0436	0,07629	0,1077	0,2238
Posti	-0,0581	0,0359	0,07382	0,1336	0,2077
Paperiteollisuus	0,0640	0,1156	0,15503	0,1895	0,2780
Muovi- ja kemian tuotteet.	-0,0411	0,0311	0,06706	0,0988	0,2103
Teleliikenneala	0,0035	0,0691	0,09529	0,1162	0,2963
Tietoliikenneala	-0,0208	0,0697	0,09507	0,1152	0,2277
Graafinen teollisuus	-0,0171	0,0977	0,12430	0,1455	0,2544
Suunn.- ja konsulttiala	-0,0262	0,0317	0,06675	0,1016	0,1881
Talonrakennusala	-0,0552	0,0126	0,04326	0,0701	0,1846
Meijerit	-0,0899	-0,0098	0,04295	0,0926	0,1441

Koulutustaso ei vaikuta millään säännönmukaisella tavalla palkkasyrjintään, mikä ilmenee taulukosta 13. Diskriminaatio on suurinta alemman korkeakouluasteen tutkinnon suorittaneiden keskuudessa ja pienintä tutkijakoulutuksen saaneilla. Havainto poikkeaa Liljan saamista tuloksista, joiden mukaan palkkadiskriminaation ja koulutustason välillä on käänteinen yhteys (1997, 27). Ilmeisesti tutkielman suurempi selittäjäjoukko vaikuttaa asiaan.

Taulukko 13. Palkkadiskriminaatio koulutustasoittain

<i>Tutkinnon taso</i>	<i>Min.</i>	<i>Q1</i>	<i>Keskiarvo</i>	<i>Q3</i>	<i>Max.</i>
Peruskoulu	-0,0899	0,0431	0,08359	0,1238	0,2780
Keskiaste	-0,0747	0,0399	0,07799	0,1174	0,2642
Alin korkea-aste	-0,0629	0,0403	0,08506	0,1250	0,2574
Alempi korkeakouluaste	-0,0619	0,0618	0,09552	0,1300	0,2647
Ylempi korkeakouluaste	-0,0565	0,0596	0,08880	0,1133	0,2763
Tutkijakoulutus	-0,0261	0,0444	0,07329	0,0973	0,2963

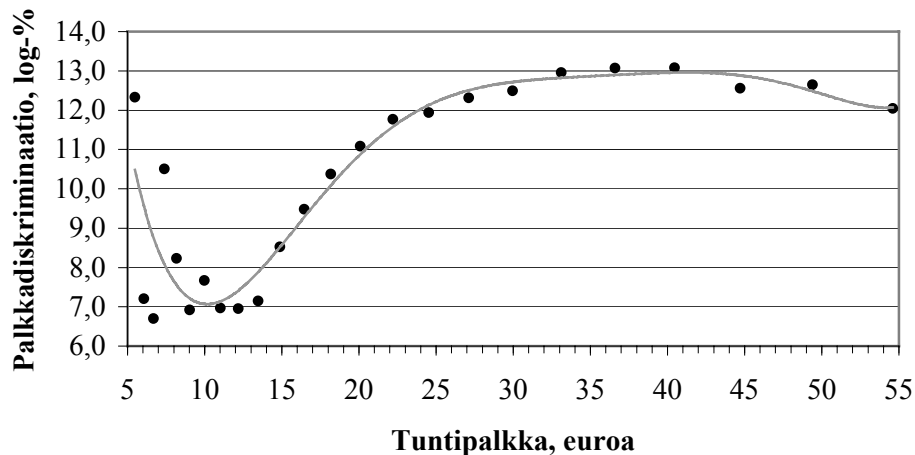
Yhteenvedon voidaan sanoa, että palkkasyrjintä vaihtelee melko paljon kaikkien tarkasteltujen ositteiden sisällä ja melkein jokaisesta löytyy muutamia negatiivisia arvoja. Sen sijaan keskimääräisissä diskriminaatioissa ei ole ositteiden välillä kovin suuria eroja ja jakaumatkin vaikuttavat melko samanlaisilta. Suurinta syr-

jintä on graafisessa ja paperiteollisuudessa ja matalinta talonrakennus- ja meijerialalla sekä teknisten toimihenkilöiden keskuudessa.

5.3 Palkkadiskriminaatio ja jatkuvat muuttujat

Tutkitaan seuraavaksi palkkadiskriminaation vaihtelua erilaisissa jatkuvien muuttujien määrittämässä ositteissa. Vartiainen esittelee hyödyllisen tavan analysoida palkkatason vaikutusta syrjintään, joten tutkielman tuloksille tehtiin samankaltainen tarkastelu (2001, 4). Koska tulojakauma on oikealle vino, tehtiin (havaitulle) palkalle ensin logaritminen luokittelu. Jokaisen toimihenkilön tuntipalkan logaritmi pyöristettiin 0,1 yksikön tarkkuudella lähimpään arvoonsa. Näin saatiin itse tuntipalkan luokitus harvenemaan jakauman loppupäässä, jossa havaintoja on vähiten. Tämän jälkeen laskettiin diskriminaatioiden keskiarvot kussakin luokassa. Kuviossa 9 esitellään näin saatu syrjinnän riippuvuus palkkatasosta. Vaaka-akselille merkityn muuntamattoman tuntipalkan luokitus siis harvenee arvojen kasvaessa. Jotta systemaattiset riippuvuudet olisi helpompi havaita, sovitettiin kuvioon tasoitettu käyrä.

Kuvio 9. Palkkadiskriminaation riippuvuus palkkatasosta



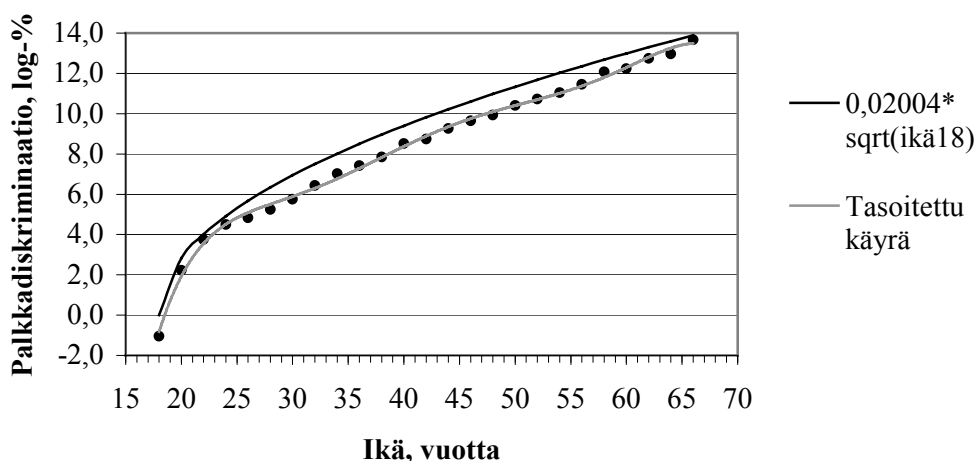
Palkkadiskriminaatio laskee ensin voimakkaasti tulotason kasvaessa, mutta kääntyy 10 euron jälkeen nousuun. Vasta aivan jakauman loppupäässä se laskee taas hiukan. Syrjintä on suurimmillaan eli melkein 13 log-prosentin tasolla 25 ja 45 euron välillä. Alhaisimmillaan eli alle kahdeksassa log-prosentissa se on 9 ja 14 euron välillä. Havainnot poikkeavat Vartiaisen saamista tuloksista, joiden mukaan teollisuuden toimihenkilöiden selittämätön palkkakaula kasvaa johdonmukaisesti ansiotason nousun myötä (2001, 60–89).

Toimihenkilöiden 13,1 euron suuruinen mediaanipalkka ja 14,4 euron suuruinen keskipalkka sijoittuvat lähelle kuvion 9 syrjinnän minimikohtaa. Palkan suuri

poikkeama keskimääräisestä ansiosta näyttää siis kasvattavan syrjintää. Erikoista on, että piirre ei riipu niinkään itse palkkatasosta, vaan nimenomaan sen ja keskiarvon välisestä erosta. Tästä voidaan epäsuorasti tehdä yllättävä johtopäätös. Jos teollisuuden toimihenkilöiden harvinaisimmat tehtävät ovat hyvin korkeasti tai matalasti palkattuja, esiintyy niissä myös suurinta palkkadiskriminaatiota. Ne olisivat siis naistoimihenkilöille erityisen haitallisia tai kääntäen miehet saisivat ylimääräistä etua niissä työskentelystä. Vastaavasti piirteiltään ja siis palkkaukseltaan tyyppillisissä tehtävissä esiintyisi vähiten syrjintää.

Kohdassa 3.2 arveltiin, että naisilla saattaisi olla jonkinlainen alkupalkkaetu miehiin nähden, mutta naisindikaattorin kautta siitä ei saatu näyttöä. Asiaa voidaan selvittää myös tutkimalla, kuinka toimihenkilöiden kohtaama palkkadiskriminaatio kehittyy iän myötä. Kuvio 10 johdettiin samalla tavalla kuin edellinenkin, mutta ikä luokiteltiin tavallisesti. Lisäksi on piirretty pelkän iän vaikutus palkkasyrjintään, jonka funktiomuoto $0,02004 * \sqrt{\text{ikä}18}$ ilmenee taulukon 9 ensimmäiseltä riviltä.

Kuvio 10. Palkkadiskriminaation riippuvuus iästä



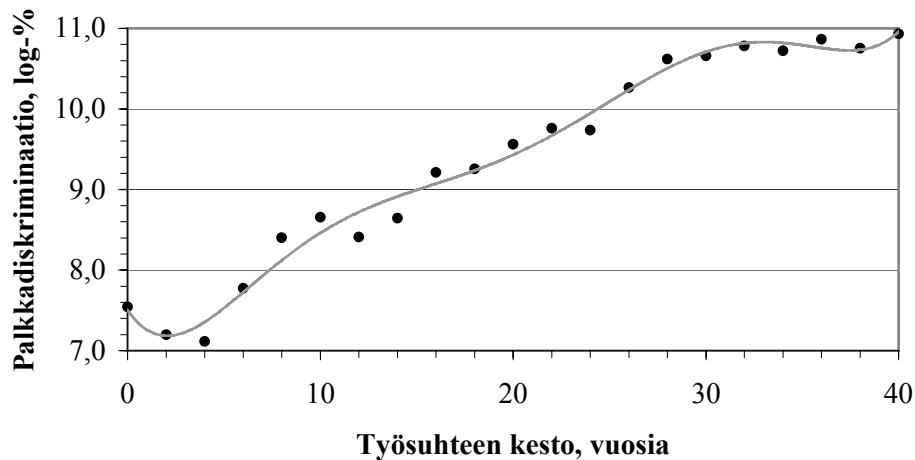
On luonnollista, että ikä nostaa voimakkaasti palkkadiskriminaatiota, sillä se on keskeisin syrjinnän lähde. Keskiarvopisteiden kautta kulkeva tasoitettu käyrä pysyy kaikkialla puhtaan ikävaikutuksen alapuolella, ja niiden pystysuora erotus on noin yksi log-prosenttiyksikkö. Voidaan päätellä, että ikätekijöiden ulkopuolelle jäävät seikat suosivat naisia tämän verran, mikä ilmenee aivan nuorimmilla naisilla juuri yhden log-prosenttiyksikön suuruisena alkupalkkaetuna. Se häviää kuitenkin jo 19 vuoden kohdalla ikäkohtelujen eron ylittäessä sen vaikutuksen. On syytä suhtautua hienoisella varauksella havaintoon sukupuolten alkupalkkojen erosta, koska iän keskiarvo on 41 vuotta ja 18-vuotiaat ovat aika kaukana aineiston painopisteestä. Siksi kohdassa 4.4 esitetty tulkinta tulee varovaisempana kyseeseen: miehet ja naiset aloittavat uransa likimain samasta tilanteesta. Diskriminaatio kasvaa ensin voimakkaasti aina 24 ikävuoteen asti. Siitä eteenpäin se

kasvaa lineaarisesti ja jokainen ikävuosi lisää syrjintää 0,25 log-prosenttiyksikköä. Eläkeiän kynnyksellä olevilla toimihenkilöillä se on melkein 14 log-%. Tämä heijastunee myös sukupuolten eläkkeisiin. Asian lähempi tarkastelu ylittäisi kuitenkin tutkielmalle asetetut tavoitteet, joten sitä ei pohdita tässä tarkemmin.

Myös Hauhio ja Lilja havaitsevat pitkittäisaineistoon perustuvassa tutkimuksessaan, että taustaominaisuuksilla selittymätön palkkaero kasvaa työuran edetessä. He arvelevat, että tilanne voi johtua naisten heikommasta halusta investoida työelämässä hyödyllisiin mutta havaitsemattomiin taitoihin. Toisena mahdollisena syynä he mainitsevat miesten nopeamman urakehityksen, jonka taustalta taas löytyisi naisten tekemän kotityön suurempi arvo. (1996, 17.) On kuitenkin vaikea sanoa, ovatko mainitut seikat palkkadiskriminaation syy vai seuraus. Kausaalisuuden suuntaa kannattaa pohtia, vaikka sitä ei voidakaan selvittää tilastollisin menetelmin. Ensimmäisen argumentin osalta voidaan todeta, että naisten halukkuus kasvattaa inhimillistä pääomaansa laskee, jos he ottavat valinnoissaan jo etukäteen huomioon investointinsa heikomman tuoton. (Vrt. Vartiainen 2001, 5–6.) Vastaavasti jälkimmäinen syy muuttuu seuraukseksi, jos naiset ajattelevat, että palkkasyrjinnän vallitessa heidän on kannattavampaa erikoistua kodinhoitoon kuin työelämään. Molemmissa tapauksissa palkkadiskriminaatio ainakin osittain ruokkisi itse itseään, kun taloudelliset toimijat mukauttavat käyttäytymisensä vallitsevaan tilanteeseen parhaiten sopivaksi. Prosessia voi vielä voimistaa se, että työn tarjonnasta sovitaan usein perheen sisällä. Havaittu käytös heijastelisi siis perheen – eikä yksilön – optimaalisia valintoja.

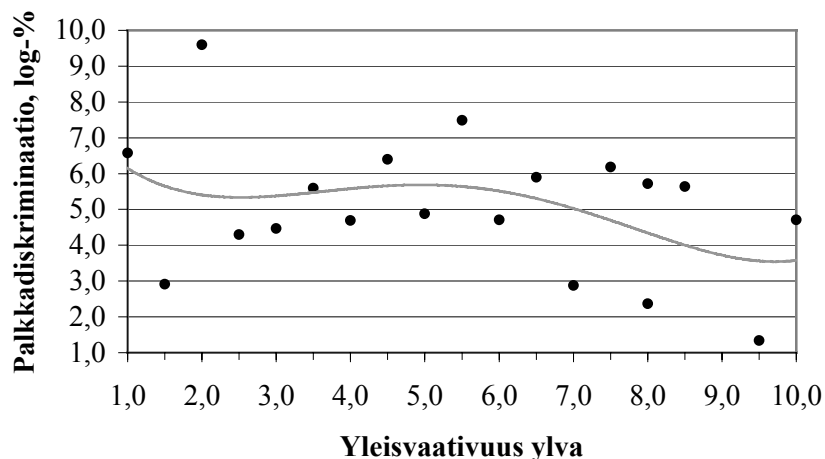
Vaikka sukupuolia palkitaan työsuhteen kestosta yhtäläisesti, nousee palkkasyrjintä sen kasvaessa. Tämä ilmenee kuviosta 11. Mitä ilmeisimmin piirre on seurausta sukupuolten erilaisesta ikäkohtelusta, sillä muuttajat ovat voimakkaasti korreloituneita. Yksi lisävuosi saman työnantajan palveluksessa johtaa tietenkin työntekijän vanhenemiseen yhdellä vuodella. Niinpä etukäteen odotettiin, että palkkadiskriminaatio kehittyisi molempien muuttajien arvojen myötä jokseenkin samalla tavalla. Ensin mainittu johtaa kuitenkin kuvion 11 valossa diskriminaation kasvuun noin 0,10 log-prosenttiyksiköllä, kun taas yksi ikävuosi johtaa kuvion 10 mukaan 0,25 log-prosenttiyksikön lisäykseen palkkasyrjinnässä. Ero on merkittävä. Diskriminaatio kyllä kasvaa tasaisesti työsuhteen pidentyessä, mutta yhden vuoden vaikutus on selvästi pienempi kuin kuvion 10 ikävaikutuksen perusteella oli odotettavissa. Ilmeisesti toimihenkilön kohtaama palkkasyrjintä kasvaa hitaammin, jos hän pysyttelee saman työnantajan palveluksessa. Siirtyminen työsuhteesta toiseen muodostaisi siis naisille ylimääräisen riskin.

Kuvio 11. Palkkadiskriminaation riippuvuus työsuhteen kestosta



Kuvion 12 valossa palkkadiskriminaatio näyttäisi yllättäen pienenevän tehtävän vaativuustason noustessa. Luonnollisesti tulos koskee vain alempia toimihenkilöitä. Haastavimmissa tehtävissä palkkasyrjintä on keskimäärin vain neljän logprosentin luokkaa, mikä on melko alhainen arvo. Hauhion ja Liljan saamien tulosten mukaan palkkasyrjintä on pienintä kaikkein kapeimmin määritellyissä vaativuusluokissa (1996, 13). Kuviossa 12 esitetty riippuvuus palautuu tähän havaintoon, jos työehtosopimuksissa kuvaillaan juuri haastavimmat tehtävät yksityiskohtaisimmin. Tämä tuntuu luontevalta oletukselta, sillä liittyihän niihin suurin palkka.

Kuvio 12. Palkkadiskriminaation riippuvuus työn vaativuudesta



5.4 Naisvaltaisuuden merkitys

On mielenkiintoinen kysymys, liittyykö yksittäisen toimihenkilön kohtaama palkkadiskriminaatio joihinkin hänen työyksikkönsä piirteisiin. TT:n aineistosta ei löydy suoraan työyksikköä kuvaavia muuttujia, mutta naisten osuus ja työnteekijöiden lukumäärä kussakin tehtävässä, yrityksessä ja sopimusallalla oli mahdollista laskea. Yhteensä toimihenkilön työpaikkaa kuvasi siis kuusi muuttujaa. Seuraavaksi muodostettiin malli PD1, jossa niillä selitettiin yksittäisen toimihenkilön kohtaamaa palkkasyrjintää. Estimoidut kertoimet olivat kyllä tilastollisesti merkitseviä, mutta ne eivät käytännössä poikenneet nolasta, eikä niitä siksi raportoida tässä⁸. Voidaan todeta, että palkkasyrjintä ei liity millään suoraviivaisella tavalla tehtävän, yrityksen tai sopimusalan kokoon tai naisvaltaisuuteen. Koev päättyy sukupuolijakauman vaikutuksen osalta samaan johtopäätökseen (1996, 49–50). Tulos on siinä mielessä odotettu, että mallin PM1 indikaattorit jo ottavat nämä piirteet hyvin pitkälle huomioon.

Työpaikkaa kuvaavia naisosuuksia ei keskistetty mallissa PD1 mitenkään. On kuitenkin ilmeistä, että työskenteleminen vaikkapa hyvin miesvaltaisen yrityksen erittäin naisvaltaisessa tehtävässä poikkeaa työskentelystä yhtä naisvaltaisessa toimessa yrityksessä, jossa on pääasiassa naisia. Tehtävän naisvaltaisuuden vaikutus palkkadiskriminaatioon voi siis vaihdella yrityksen sukupuolijakauman mukaan. Vastaava argumentointi voidaan toistaa yritysten ja toimialojen tapauksessa. Asian tutkimiseksi laskettiin ensin naisten osuus jokaisessa työsolussa (engl. job-cell), joka määritellään yrityksen sisäisenä tehtävänä⁹ (Groshen 1991, 458). Alle 200 hengen yrityksiä käsiteltiin yhtenä yksikkönä, jonka sopimusalaksi valittiin sen moodi. Näin saatiin luotua hierarkkinen rakenne: jokainen toimihenkilö kuuluu johonkin työsoluun, joka edelleen on osa jotain tiettyä yritystä. Kukin yritys taas toimii vain yhdellä alalla. Seuraavaksi siirryttiin tarkastelemaan *poikkeamia* hierarkian ylemmältä tasolta: jokaisen yrityksen naisosuudesta vähennettiin sopimusalan naisosuus, ja jokaisen työsolun naisosuudesta vähennettiin yrityksen naisosuus. Saadut arvot kerrottiin sadalla, jotta mittayksiköksi olisi saatu prosentteja. Lopuksi estimoitiin malli PD2, jossa selitettiin johdetuilla kolmella muuttujalla yksittäisen toimihenkilön kohtaamaa log-prosenteissa ilmaistua palkkasyrjintää¹⁰. Tulokset raportoidaan taulukossa 14.

⁸ Ratkaisuun päädyttiin osittain myös siksi, että mallin PD1 kertoimien hajonnat saattavat olla aliarvioituja, koska regression residuaalit eivät välttämättä täytä kaikkia Gauss-Markov-ehtoja (Dougherty 1992, 80–82). Aineiston ryhmittäinen rakenne yhdistettynä ryhmää kuvaavien muuttujien käyttöön selittäjinä saattaa johtaa residuaalien korreloituneisuuteen ryhmien sisällä. Tällä voi olla estimoinnin tarkkuutta ylliarvioiva vaikutus. Itse kerroinestimaatteihin ilmiö ei sen sijaan vaikuta. (Kloek 1981, 205–207 ja Moulton 1990, 334–338.)

⁹ Aivan tutkielman teon alkuvaiheessa kokeiltiin palkkamallia, jossa oli selittäjinä mm. työsoluindikaattorit. Niitä oli kuitenkin niin paljon, että malli ei ollut estimoitavissa ja siksi niiden käytöstä luovuttiin.

¹⁰ Alaviitteessä 8 esiin tuodut seikat koskevat myös mallia PD2.

Taulukko 14. Malli PD2

<i>Muuttuja</i>	<i>Kerroin</i>	<i>t-arvo</i>	<i>P(ChiSq)</i>
vakio	8,50011	624,99	<0,0001
sopimusalan naisvaltaisuus	0,01622	10,28	<0,0001
yrittäjän naisvaltaisuus, poikkeama sopimusala-	0,04639	24,06	<0,0001
työolosuhteiden naisvaltaisuus, poikkeama yri-	0,00672	15,61	<0,0001
työstä			

Mallin PD2 selitysaste on hyvin matala, vain 0,62 prosenttia. Tulokset ovat kuitenkin mielenkiintoisia. Kaikki selittäjät kasvattavat palkkadiiskriminaatiota, mikä tarkoittaa, että toimihenkilön kohtaama syrjintä kasvaa, jos hänen työyksikkönsä on selvästi ympäristöään naisvaltaisempi. Vastaavasti diiskriminaatio pienenee, jos toimihenkilön työyksikkö on ympäristöään miesvaltaisempi. Tämä pätee riippumatta siitä, tarkastellaanko sopimusalaa, yritystä vai työolosuhteita. Selvästi voimakkain vaikutus on yrityksen sukupuolijakauman poikkeamalla sopimusalan sukupuolijakaumasta. Keskistetyt naisosuuden nousu kymmenellä prosenttiyksiköllä johtaa toimihenkilön kohtaaman palkkadiiskriminaation kasvuun keskimäärin noin puolella log-prosenttiyksiköllä. Naisille erityisen epäedullisia (tai miehille edullisia) ovat siis miesvaltaisten alojen naisvaltaiset yritykset. Yllättäen työolosuhteiden naisvaltaisuuden merkitys pysyy myös mallissa PD2 vähäisenä.

5.5 Keillä palkkasyrjintä on erityisen suurta tai pientä?

Yksittäisen toimihenkilön kohtaama palkkadiiskriminaatio vaihtelee aineistossa melko paljon. Matalin arvo on -8,99 ja korkein 29,63 log-prosenttia. Aiempien analyysien avulla saatiin kuva palkkasyrjintää voimakkaimmin määrittävistä tekijöistä, mutta on myös hyödyllistä tarkastella suoraan niitä toimihenkilöitä, joiden kohtaama diiskriminaatio on poikkeuksellisen suurta tai pientä.

5011 toimihenkilöllä palkkasyrjintä on negatiivinen tai nolla. Joukko koostuu naisista, joilla on palkkaetua samassa työtilanteessa oleviin miehiin nähden ja miehistä, joiden tilanne on päinvastainen. Ryhmä on melkoisen pieni aineiston kokoon verrattuna, mistä voidaan päätellä, että naisten palkkaetu tai miesten palkkahaitta ovat suhteellisen harvinaisia ilmiöitä. Kun tätä joukkoa tutkittiin tarkemmin, havaittiin, että sen sopimusala- ja sukupuolijakauma ovat pitkälle yhteneviä muun aineiston kanssa. Sen sijaan selkeimmin ryhmä eroaa muusta aineistosta ikäkeskiarvon osalta. Toimihenkilöt ovat kyseisessä joukossa selvästi muuta aineistoa nuorempia, keskimäärin vain 32,4 vuotta vanhoja. He ovat myös hiukan muita aineiston toimihenkilöitä heikommin koulutettuja. Ero on noin yksi koulutusvuosi. Lisäksi he ovat olleet nykyisen työnantajansa palveluksessa keskimääräistä lyhyemmän ajan. Ryhmä muut toimihenkilöt on 60 prosentin osuudellaan voimakkaasti ylliedustuttuna tässä 5000 hengen osajoukossa, kun taas

ylempiä toimihenkilöitä löytyy siitä vain yksi prosentti. Erityisen suurta eli 20 log-% tai yli palkkadiskriminaatio on 2421 toimihenkilöllä. Ryhmä koostuu pääasiassa vanhemmista ylemmistä toimihenkilöistä, joiden meneillään oleva työsuhte on kestänyt hyvin pitkään, keskimäärin yli 20 vuotta. Paperiteollisuus on 58 prosentin osuudellaan selvästi yliedustettuna tässä ryhmässä, kun taas metalliteollisuuden toimihenkilöitä on vain 5,5 prosenttia. Naisia on 28 prosenttia, mikä on hiukan muuta aineistoa alempi luku.

Tulokset ovat odotettuja. Palkkadiskriminaatiofunktion kertoimista (taulukko 9) nähdään, että ikä vaikuttaa voimakkaimmin toimihenkilön kohtaamaan syrjintään ja että paperiteollisuuteen kuulumisen nostaa palkkasyrjintää. Havainnot poikkeavat osittain Koevin saamista tuloksista. Hän päätyy johtopäätökseen, jonka mukaan teollisuudessa ei juurikaan ole työtilanteita, joissa naissukupuolesta olisi palkan kannalta etua (tai mieheydestä haittaa). (Koev 1996, 50–51.) Tämän kohdan tarkastelujen valossa näyttää siltä, että kyseinen ilmiö on hyvin harvinainen, mutta ei kuitenkaan olematon. Yli 5000 teollisuuden toimihenkilöä tekee työtä, jossa palkkasyrjintä arvioidaan negatiiviseksi.

5.6 Miesten ja naisten erilaisesta kohtelusta

Edellisissä tarkasteluissa pohdittiin palkkadiskriminaation jakautumista teollisuudessa ja sen kehittymistä jatkuvien selittäjien saamien arvojen muuttuessa. Niissä miesten ja naisten erilainen kohtelu työmarkkinoilla tiivistettiin yhteen tunnuslukuun, eli kohdattuun palkkasyrjintään. Asetelma voidaan ja kannattaakin kääntää myös toisin päin. Kysytään, kuinka paljon osaavampi tai kokeneempi naisen tulee olla, jotta hän saavuttaisi palkan, joka miehellä arvioidaan olevan hänen työtilannettaan vastaavassa asemassa. Tällaista kysymyksenasettelua on pidetty esillä varsin paljon julkisessa keskustelussa, ja palkkasyrjinnän arviointiin käytetty menetelmä soveltuu myös sen pohtimiseen.

Keskimääräiseksi palkkadiskriminaatioksi arvioitiin 8,5 log-%. Ajatellaan, että kurotaan tämä miesten palkkauksellinen etumatka umpeen lisäämällä naisille jotain ansioita kasvattavaa ominaisuutta – vaikkapa iän tuomaa työkokemusta. Tämä on luonnollinen valinta, koska se on tärkein yksittäinen palkkasyrjintää kasvattava tekijä. Kuvioista 6 arvioituna ikä nostaa logaritmista palkkaa työuran aikana keskimäärin 0,013 yksikköä vuodessa. Tästä saadaan laskettua, että naisten tulisi olla 6,5 vuotta miehiä vanhempia, jotta he muuten identtisessä työssä saisivat keskimäärin samaa palkkaa. Koska työuran alussa ikä kasvattaa palkkoja nopeammin, riittää nuoremmilla naisilla pienempi iän lisäys miesten palkkaedun umpeen kuromiseksi. Vastaavasti vanhemmilla naistoimihenkilöillä tähän tarvitaan useampia vuosia ylimääräistä työkokemusta. Sama analyysi voidaan toistaa koulutusvuosien kohdalla. Kuvion 7 perusteella yksi koulutusvuosi kasvattaa logaritmista palkkaa keskimäärin 0,025 yksiköllä ja naiset siis tarvitsisivat 3,4

vuotta miehiä enemmän koulutusta, jotta syrjivä palkkaetu saataisiin kurotuksi umpeen. Tämä vastaa kokonaista keskiasteen tutkintoa.

Tulokset ovat hyvin mielenkiintoisia varsinkin iän osalta. Melko yleisesti nimittäin väitetään, että palkkadiskriminaatio johtuisi naisten miehiä katkonaisemmasta työurasta, jonka taustalta taas löytyisivät äitiyslomat (Asplund 1996, 10). Suomessa naiset hankkivat keskimäärin kaksi lasta ja miesten 6,5 vuoden ikäetu naisiin nähden on hämmästyttävän lähellä sitä ajanjaksoa, jonka voitaisiin luontevasti arvioida kuluvan kahden pienen lapsen hoitamiseen – ikäetuhan olisi aivan hyvin voinut osoittautua vaikkapa yhden tai 20 vuoden suuruiseksi. Havainnon merkitystä pohditaankin tarkemmin luvussa 7. Tietenkin on huomioitava, että osoitettu yhteys on puhtaasti deskriptiivinen tulos ja siis vailla kausaalitulkintaa. Poikkileikkausaineistoon perustuvien tilastollisten tarkastelujen ei voida mitenkään saada selville, mikä on aiheuttanut naisten jäämisen yli kuudella vuodella jälkeen miesten palkkakehityksestä. Edelleen laskelmat ovat ainakin osittain hypoteettisia, ja siksi niitä tulee pitää ainoastaan suuntaa-antavina. Yhtä kaikki, tulokset ovat erittäin kiinnostavia ja havainnollistavia ja ansaitsevat siksi tulla raportoiduiksi. Koska pelkät palkkadiskriminaatioarviot ovat melko abstrakteja, konkretisoivat tällaiset tarkastelut oivalla tavalla sukupuolten erilaista kohtelua työmarkkinoilla. Niiden avulla saadaan kuva siitä, mitä palkkasyrjintä käytännössä tarkoittaa. Onkin yllättävää, ettei vastaavia laskelmia löydy muista tutkielman lähteinä käytetyistä palkkaerotutkimuksista.

6 Tulosten luotettavuuden ja oikeellisuuden arviointi

6.1 Palkkamalli GPM1 ja palkkadiskriminaation luottamusväli

Mallin PM1 perusteella arvioitiin palkkadiskriminaatioksi toimihenkilöiden tyypillisessä tehtävässä 8,53 log-%. On hyödyllistä arvioida käytetyn menetelmän tilastollista luotettavuutta laskemalla arviolle 99 prosentin luottamusväli. Tämä on melko työläs tehtävä, jos se tehdään suoraan malliin PM1 perustuen. Siksi hyödynnettiin palkkasyrjinnän ja hyödykkeiden laatuvaikoidun hinnanmuutoksen tutkimisen yhtenevyyttä. Palkkadiskriminaation laskeminen vastaa nimittäin teknisesti tarkalleen miesten ja naisten välisen palkkaeron puhdistamista sukupuolten työtehtävien ja ominaisuuksien välisistä eroista hedonisella regressiolla, jota käytetään tuotteiden laadun vakioimiseen tutkittaessa niiden hinnanmuutoksia. Niinpä tässä sovellettiin yhtä laatuvaikoinnin menetelmistä, Griliches-regressiota, joka on eräänlainen hyödyllinen virhespesifikaatio. Siinä tuotteiden hintoja eri ajankohtina selitetään niiden ominaisuuksilla, mutta hintayhtälöiden kertoimien ei anneta muuttua ajassa. Griliches-tyyppisissä hintayhtälöissä sallitaan ainoastaan regressiovakion vaihtelu ajassa. (Vartia – Koskimäki 2001.) Palkkadiskriminaation tapauksessa täysin analoginen on malli, jossa miesten ja naisten palkkayhtälöt pakotetaan vakiota lukuunottamatta identtisiksi. Tätä tekniikkaa mukailien muodostettiin malli GPM1 poistamalla mallista PM1 kaikki muotoa $n \cdot xxx$ olevat interaktiot ja lisäämällä siihen naisindikaattori. Muilta osin GPM1 on siis yhtenevä PM1:n kanssa.

Kun GPM1 estimoitiin, saatiin naisindikaattorin kertoimeksi -0,0832. Sen antama palkkadiskriminaatioarvio on siis 8,32 log-prosenttia, ja se on yllättävän lähellä mallin PM1 avulla laskettua arviota. GPM1:n selitysaste on 0,7735 eli toisin sanoen puoli prosenttiyksikköä PM1:tä matalampi. Naisindikaattorin hajonnasta mallissa GPM1 laskettiin luottamusväli, jota edelleen sovellettiin mallin PM1 palkkadiskriminaatioarvion ympärille. Näin saatiin palkkasyrjinnän 99 prosentin luottamusväliksi tyypillisessä toimihenkilötehtävässä 8,23–8,83 log-%. Väli on melko kapea, mikä osoittaa, että käytetty menetelmä on puhtaasti tilastollisessa mielessä hyvin tarkka. Tämä johtunee erittäin suuresta aineistosta.

6.2 Mahdollisten virhespesifikaatioiden etsiminen malleilla PM2 - PM15

Vaikka PM1 on hyvin joustava, se ei salli kertoimien vapaata vaihtelua kaikissa mahdollisissa aineiston ositteissa. Niitä on hyvin paljon, eikä ollut tarkoituksenmukaista sirpaloida aineistoa liian pieniin osiin. Tarvittavat interaktiot (ks. taulukko 6) pääteltiin aiemman tutkimuksen perusteella, ja niiden avulla sallittiin jatkuvien muuttujien kertoimien vapaa vaihtelu sukupuolen ja toimihenkilöryhmän mukaan. Toisaalta kohdassa 2.4 esitettiin, kuinka segregaatian ja aineiston

heterogeenisuuden välinen vuorovaikutus voi vaikeuttaa palkkadiskriminaation tutkimista. Jos vaikkapa miehet ovat hakeutuneet yrityksiin, joissa iästä palkitaan keskimääräistä paremmin, näyttää heillä aggregaattitason regressiossa olevan naisia suurempi ikäkerroin. Tällöin palkkasyrjintä yliarvioidaan, jos ikäkertoimen ei anneta vaihdella yrityksittäin. Käytettävä malli saattaa myös olla sillä tavalla liian jäykkä, että ositteiden sisäiset palkkatasoerot jäävät havaitsematta. Tutkimatta on mahdotonta sanoa, missä määrin – jos ollenkaan – PM1:llä tehdyt johtopäätökset kärsivät näistä seikoista. Seuraavaksi esitellään menetelmä, jota käytettiin tulosten oikeellisuuden selvittämisessä.

Mallissa PM1 on neljä jatkuvaa muuttujaa (ks. taulukko 5), ja aineisto voidaan jakaa sopimusaloihin, yrityksiin ja tehtäviin. Jokaisen jatkuvan muuttujan kerroin saattaa poiketa jossain mainitussa aineiston ositteessa päättelyä vaarantavalla tavalla. Siksi muodostettiin ja estimoitiin palkkamallit, joissa interaktioita lisämällä sallittiin kerrallaan aina yhden muuttujan kertoimien vapaa vaihtelu joko sopimusaloilla, yrityksissä tai tehtävissä. On syytä huomioida, että ei ole niinkään tärkeää, ovatko lisätyt interaktiot tilastollisesti merkitseviä vai eivät. Sen sijaan ratkaisevaa on, muuttaako niiden lisääminen malliin PM1 johtopäätöksiä eli sen avulla laskettua palkkadiskriminaatiota. Siksi ei kiinnitetty huomiota yksittäisiin kertoimiin, vaan jokaisen edellä kuvatulla tavalla muodostetun mallin antamaan palkkasyrjintäarvioon, jota verrattiin mallin PM1 antamaan lukemaan. Päättely perustettiin niiden väliseen erotukseen. Esitellään seuraavaksi esimerkkinä, kuinka tällä menetelmällä tutkittiin, olisiko mallissa PM1 pitänyt sallia muuttujan $\sqrt{\text{ikä}18}$ kertoimen vapaa vaihtelu sopimusaloittain.

Ensimmäinen muodostettu malli on nimeltään PM2, ja se saatiin lisäämällä PM1:een sopimusalaindikaattorien ja muuttujan $\sqrt{\text{ikä}18}$ väliset interaktiot. Näin sallittiin vapaa iästä palkitseminen kaikilla sopimusaloilla. Kun malli estimoitiin, saatiin palkkadiskriminaatiofunktiolle uudet taulukon 9 arvoista hieman poikkeavat kertoimet. Niiden perusteella laskettiin uudelleen palkkadiskriminaatio keskimääräisessä toimihenkilötehtävässä, jota verrattiin vastaavaan mallin PM1 avulla saatuun syrjintäarvioon. Jos nyt vaikkapa miehet olisivat hakeutuneet iästä paremmin palkitseville aloille, olisi muuttujan $\sqrt{\text{ikä}18}$ kerroin PDF:ssa laskenut voimakkaasti taulukon 9 arvosta. Tällöin mallien PM1 ja PM2 antamat syrjintäluvut olisivat poikenneet selvästi toisistaan, mitä olisi voitu pitää merkkinä siitä, että PM1 on liian jäykkä malli. Jos taas luvut olivat lähellä toisiaan, voitiin päätellä, että tutkitun palkanmuodostuksen piirteen – mallin PM2 tapauksessa sopimusaloittain vaihtelevan ikäkertoimen – puuttuminen mallista PM1 on harmiton yksinkertaistus. Tällaista tekniikkaa sovellettiin kaikkiaan 14 kertaa. Menettely kannattaa tulkita sarjaksi äärimmäisen sensitiivisiä mallille PM1 tehtäviä testejä. Jos siitä nimittäin puuttuu joitain palkkadiskriminaation arvioinnin kannalta olennaisia toimihenkilöiden palkanmuodostuksen piirteitä, paljastuu tarve niihin tällä tekniikalla aivan varmasti.

Taulukossa 15 esitellään mallit PM2 - PM5, joissa on sallittu jatkuvien muuttujien kertoimien vaihtelu sopimusaloittain. Vertailun vuoksi ensimmäiseltä riviltä ilmenevät mallin PM1 tiedot. Kukin malli on saatu lisäämällä PM1:een joukko interaktioita, joiden tyyppi esitetään toisessa sarakkeessa. Sarake $PDF(\bar{\mathbf{x}})$ ilmaisee jokaisen näin saadun mallin antaman palkkadiskriminaatioarvion tyypillisessä toimihenkilötehtävässä. Neljännessä sarakkeessa raportoidaan kunkin mallin selitysaste.

Taulukko 15. Sopimusaloittain vaihtelevien kertoimien merkitys

Mallin nimi	PM1:een lisättyjen interaktioiden tyyppi	$PDF(\bar{\mathbf{x}})$	R^2
PM1	-	0,0853	0,7784
PM2	sopimusalat*sqrt(ikä18)	0,0851	0,7788
PM3	sopimusalat*sqrt(tysk)	0,0850	0,7791
PM4	sopimusalat*kv8	0,0858	0,7812
PM5	sopimusalat*ylva	0,0846	0,7827

Yksikään malleista PM2 - PM5 ei tuota PM1:stä olennaisesti poikkeavaa johtopäätöstä palkkasyrjinnästä. Huomionarvoista on, että mallissa PM5 sallitaan käytännössä jokaiselle liitteen 1 vaativuusasteikolle oma kerroin; teknisten 6- ja konttoritoimihenkilöiden 12-portaisille asteikoillehan jo sallitaan PM1:ssä erilaiset kertoimet. PM5 ja PM1 arvioivat kumpikin palkkadiskriminaation tyypillisessä toimihenkilötehtävässä kahdeksan ja puolen log-prosentin suuruiseksi. Tämä osoittaa, että ylvan johtamisessa tehty vaativuusasteikoiden skaalaus (11) on harmiton yksinkertaistus. Se venyttää lyhyet ja kutistaa moniportaiset asteikot niin tehokkaasti, että ylva voidaan käyttää analyysissä sellaisenaan johtopäätösten siitä muuttumatta.

Taulukossa 16 esitetään tulokset, jotka saatiin, kun mahdollistettiin jatkuvien muuttujien kertoimien vapaa vaihtelu yli 200 hengen yrityksissä. Koska mallissa PM1 ei sallittu miesten ja naisten palkkatasoeroja yritysten sisällä, selvitettiin lisäksi kyseisen piirteen merkitys mallin PM10 avulla. Se saatiin lisäämällä PM1:een nais- ja yritysindikaattorien interaktiot, jotka mahdollistivat sukupuolten erilaiset palkkatasot yritysten sisällä. Tällöin tietysti PDF:n argumenttien määrä kasvoi 22:sta kyseisten interaktioiden verran ja mallin PM10 palkkasyrjintäarvio tyypillisessä toimihenkilötehtävässä laskettiin näin saadulla pidemmällä PDF:lla. Jälleen havaitaan, että mallien PM6 - PM10 antamat arviot palkkadiskriminaatiosta tyypillisessä toimihenkilötehtävässä ovat yhteneviä mallin PM1 tulosten kanssa.

Taulukko 16. Yrityksittäin vaihtelevien kertoimien merkitys

Mallin nimi	PM1:een lisättyjen interaktioiden tyyppi	PDF($\bar{\mathbf{x}}$)	R ²
PM6	yrietykset*sqrt(ikä18)	0,0851	0,7799
PM7	yrietykset*sqrt(tysk)	0,0849	0,7808
PM8	yrietykset*kv8	0,0858	0,7814
PM9	yrietykset*ylva	0,0846	0,7825
PM10	n*yrietykset	0,0854	0,7802

Taulukon 16 tuloksista seuraa eräs huomio, joka liittyy käytyyn palkkadiskriminaation tutkimista koskevaan keskusteluun. Pasanen osoitti, ettei palkkasyrjinnästä muutu, vaikka palkkamalliin lisätään yritysindikaattorit (1995). Tämän perusteella Vartia argumentoi, että yritystason tekijät voidaan palkkadiskriminaatiota tutkittaessa sivuuttaa (1995, 19–20). Yritystä kuvaavien muuttujien osalta näin epäilemättä onkin. Kaikki niiden lineaarikombinaatiot johtavat vain erisuuruisiin palkkatasoefekteihin, jotka yritysindikaattorit tehokkaasti ottavat huomioon. Jos sen sijaan selittäjien kertoimet vaihtelevat yrityksittäin, on tilanne toinen. Potentiaalisestihan ilmiöllä voi segregaaation vuoksi olla tuloksia huomattavasti muuttava vaikutus, mikä hyvin ilmenee kohdan 2.4 esimerkeistä. Tätä ei saada selville yritysindikaattoreiden avulla, koska ne kontrolloivat ainoastaan erilaisia palkkatasoja. Mallien PM6 - PM10 avulla tehty analyysi osoittaa, että myös yrityksittäin vaihtelevat kertoimet voidaan sivuuttaa palkkasyrjintää tutkittaessa.

Lopuksi tutkittiin mallien PM11 - PM15 avulla tehtävittäin vaihtelevien kertoimien merkitystä. Tulokset raportoidaan taulukossa 17, josta havaitaan, etteivät mallien PM11, PM12 ja PM14 avulla saadut tulokset juurikaan poikkea mallista PM1. Sen sijaan mallien PM13 ja PM15 antamat palkkasyrjintäarviot ovat jonkin verran pienempiä kuin vastaava PM1:een perustuva luku. Erityisesti muotoa n *tehtävät olevien interaktioiden lisääminen malliin PM1 johtaa palkkadiskriminaation laskuun 0,7 log-prosenttiyksiköllä. Tämä johtuu siitä, että mallin PM15 tuottamassa PDF:ssa muuttujan $\sqrt{\text{ikä18}}$ kerroin laskee taulukon 9 arvosta 0,02004 lukemaan 0,01880. Koska miesten ja naisten erilainen ikäkohtelu on tärkein yksittäinen palkkasyrjinnän lähde, laskee arvio diskriminaatiosta vastaavasti. Voidaan päätellä, että PM1:n avulla ei havaita kaikkia tehtäväluokkien sisällä vallitsevia sukupuolten välisiä palkkatasoeroja, vaan se hienokseltaan yliarvioi palkkasyrjinnän. Havainnon merkitystä pohditaan tarkemmin seuraavassa kohdassa, jossa tehdään yhteenveto tämän luvun tuloksista.

Taulukko 17. Tehtävittäin vaihtelevien kertoimien merkitys

<i>Mallin nimi</i>	<i>PM1:een lisättyjen interaktioiden tyyppi</i>	<i>PDF(\bar{x})</i>	<i>R²</i>
PM11	tehtävät*sqrt(ikä18)	0,0847	0,7793
PM12	tehtävät*sqrt(tysk)	0,0845	0,7797
PM13	tehtävät*kv8	0,0827	0,7881
PM14	tehtävät*ylva	0,0847	0,7815
PM15	n*tehtävät	0,0783	0,7800

6.3 Yhteenveto luvusta 6

Kun luvun 5 arviolle palkkadiskriminaatiosta laskettiin luottamusväli, hyödynnettiin mallia GPM1. Huomionarvoista on, että sen avulla voidaan tehdä oikeasuuntaista päättelyä palkkasyrjinnästä, vaikka se on puhtaasti tilastollisin kriteerein arvioituna ilmiselvästi virheellinen. Sehän muodostettiin poistamalla PM1:stä 22 tilastollisesti erittäin merkitsevää muuttujaa ja lisäämällä siihen yksi alun perin täysin merkityksetön selittäjä. Toki PM1 on GPM1:tä selvästi tarkempi ja siksi informatiivisempi kuvaus palkkasyrjinnän taustalta löytyvistä mekanismeista. Silti on erikoista, että näin rajusti yksinkertaistettu malli tuottaa käytännössä saman johtopäätöksen kuin sitä selvästi monimutkaisempi PM1. Lisäksi selitysaste laski suhteellisen vähän siirryttäessä PM1:stä GPM1:een. Poistetuilla interaktioilla ei juurikaan ole itsenäistä kykyä selittää toimihenkilöiden palkkaa tai palkkasyrjintää tavalla, johon naisindikaattori ei pystyisi. Osittain nämä havainnot palautuvat alun pohdintaan tilastollisen ja asiallisen merkitsevyyden erosta. Toisaalta tuloksella on itseisarvoa myös tilastotieteen teorian kannalta, koska se osoittaa, että sovellettu Griliches-tyyppinen hedoninen regressio on tämän tutkielman tapauksessa toimiva menetelmä¹¹ ja siis hyödyllinen virhespesifikaatio. Asiaa ei kuitenkaan pohdita tässä tarkemmin, sillä se ei kuulu tutkielman piiriin.

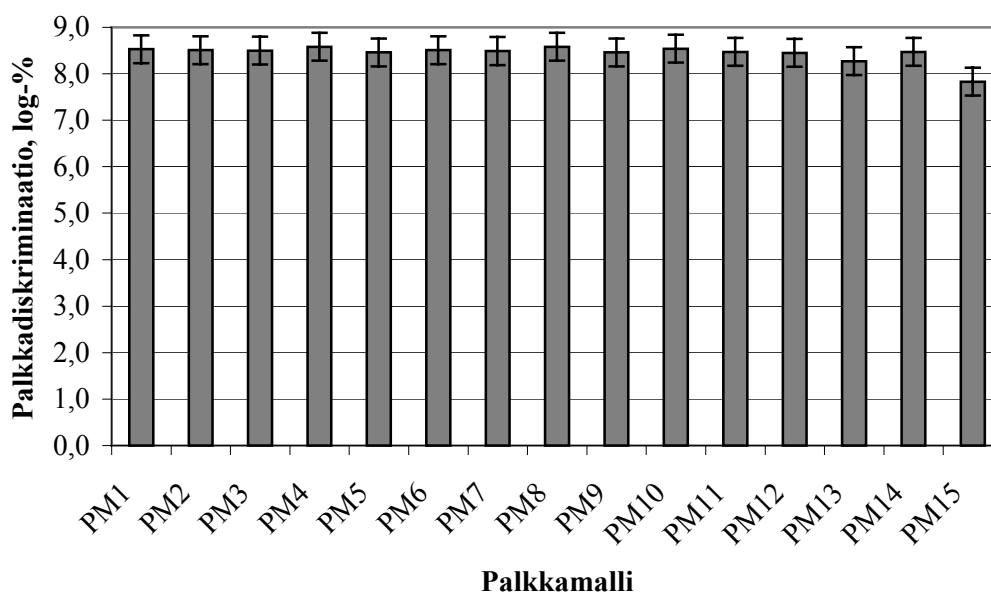
GPM1:n avulla saatu 99 prosentin luottamusväli osoitti, että käytetyn menetelmän tilastollinen osumatarkkuus on hyvä. Kuitenkin yhdessä mallille PM1 tehdystä 14:stä testistä havaittiin, että olisi ollut perusteltua tarkentaa sen rakenteellista osaa tehtävatasolla. Muotoa n*tehtävät olevien interaktioiden lisääminen malliin johti nimittäin sen antaman palkkadiskriminaatioarvion laskuun 7,83 log-prosenttiin. Miten tähän tulokseen kuuluisi suhtautua? Ainakaan mallin PM15 tuomaa informaatiota ei voida sivuuttaa. Palkkasyrjintäarvion muutos 0,7 log-prosenttiyksiköllä on sen verran suuri, että periaatteessa kaikki lukujen 4 ja 5 estimoinnit ja niihin perustuvat laskelmat olisi pitänyt tehdä uudelleen ottamalla

¹¹ Vartia ja Koskimäki painottavat, että Griliches –tyyppisen hedonisen regressio toimivuutta ei voida päätellä etukäteen, vaan sen antamia tuloksia on aina verrattava joustavampaan malliin (Vartia – Koskimäki 2001).

lähtökohdaksi malli, jossa olisi kyseisten interaktioiden käytöllä sallittu miesten ja naisten erilaiset palkkatasot tehtävien sisällä. Etukäteen tätä ei voitu tietää, eikä tutkielman tekoon käytettävissä ollut aika valitettavasti mahdollistanut tällaista ratkaisua. Niinpä saatuihin tuloksiin on suhtauduttava siten, että ne hiukan yliarvioivat todellista palkkadiskriminaatiota.

Toisaalta on otettava huomioon, että PM1 osoittautui tuloksiltaan hyvin lähellä oikeaa olevaksi malliksi, jos PM15 sellaiseksi halutaan tulkita. Käsitystä palkkadiskriminaation suuruusluokasta ei tarvitse muuttaa, sillä se pysyi samana malleissa PM1 - PM15. Tämä ilmenee hyvin kuviosta 13, jossa esitetään yhteenvedonomaaisesti mallien PM1 - PM15 antamat arviot palkkasyrjinnästä tyypillisessä toimihenkilötehtävässä. Niille on myös laskettu GPM1:n avulla 99 prosentin luottamusvälit.

Kuvio 13. Mallien PM1 - PM15 palkkadiskriminaatioarviot ja niiden luottamusvälit



On oikeastaan hyvin yllättävää, etteivät arviot palkkasyrjinnästä muuttuneet tämän enempää segregaaion ja heterogeisuuden vuorovaikutuksen tutkimisen myötä. Potentiaalisestihan ilmiö olisi voinut olla kuinka vakava ongelma tahansa, kuten kohdan 2.4 esimerkeistä hyvin ilmenee. Tulokset olisivat voineet muuttua sen huomioimisesta täydellisesti jopa etumerkkiä myöten. Täytyy myös pitää mielessä, että PM1:n testaamiseen käytetty menetelmä on äärimmäisen sensitiivinen. Se ei anna mitään mallin puutteita anteeksi, vaan paljastaa ne armotta, ja niin sen pitääkin tehdä. Näihin seikkoihin nähden PM1 selvisi 14 testin sarjasta suorastaan erinomaisesti – arvio palkkasyrjinnästä tyypillisessä toimihenkilötehtävässä laski enimmilläänkin vain 0,7 log-prosenttiyksikköä, ja 12:ssa testissä ei

ilmennyt mitään tarvetta korjata mallia. Edelleen on otettava huomioon, että epä-kokeelliseen aineistoon perustuvassa päättelyssä tulokset ovat aina epävarmempia kuin pelkät suoraviivaiset tilastolliset tunnusluvut välittömästi antavat ymmärtää. Niinpä tutkielman luonteva johtopäätös on, että palkkadiskriminaatio toimihenkilöiden keskimääräisessä työtilanteessa on noin 8 log-%. PM1:n avulla saatuja syrjintää koskevia deskriptiivisiä tuloksia ei tarvinne muuttaa lainkaan.

7 Johtopäätökset

Luvussa 1 esitettiin kolme keskeistä teollisuussektorin toimihenkilöiden palkkadiskriminaatiota koskevaa kysymystä. Niihin haettiin vastausta muodostamalla inhimillisen pääoman teoriaan perustuva yksityiskohtainen palkanmuodostusta kuvaava tilastollinen malli, josta edelleen johdettiin palkkadiskriminaatiofunktio. Sen avulla pääteltiin tekijät, joista palkkasyrjintä keskeisesti riippuu ja laskettiin diskriminaation suuruus teollisuuden eri osissa. Lopuksi tutkittiin saatujen tulosten oikeellisuutta ja havaittiin, että niitä on syytä hiukan tarkistaa alaspäin, mutta että niiden suuruusluokka on oikea.

Teollisuussektorin toimihenkilöillä esiintyy palkkaeroja samaa työtä tekevien miesten ja naisten välillä. Suuruudeltaan tämä palkkasyrjintä on tyypillisessä toimihenkilötehtävässä noin 8 log-%¹². Loppuosa eli noin kaksi kolmannesta sukupuolten keskipalkkojen välisestä erosta johtuu miesten ja naisten erilaisista henkilökohtaisista ominaisuuksista ja eroista heidän työnsä piirteissä. Arvio palkkadiskriminaatiosta on aiempia palkkaerotutkimuksia selvästi alempi. Yksittäisen toimihenkilön kohtaama palkkasyrjintä vaihtelee melko paljon, ja yli 5000 henkilöllä se on negatiivinen. Kahdella ja puolella tuhannella toimihenkilöllä palkkadiskriminaatio taas ylittää 20 log-prosenttia. Erityisen suurta syrjintä on ylempien toimihenkilöiden keskuudessa ja paperi- ja graafisessa teollisuudessa. Poikkeuksellisen pientä se taas on teknisten toimihenkilöiden parissa ja talonrakennus- ja meijerialalla.

Yllättäen toimihenkilön kohtaama palkkasyrjintä keskimäärin kasvaa, jos hänen työyksikkönsä on selvästi ympäristöään naisvaltaisempi. Tämä pätee siitä riippumatta, verrataanko sopimusalaa aineistoon, yritystä sopimusalaan vai työsolua yritykseen; suurin vaikutus on yrityksen naisvaltaisuuden poikkeamalla sopimusalan naisvaltaisuudesta. Toinen yllättävä havainto on, että toimihenkilön palkan *poikkeama* keskipalkasta vaikuttaa diskriminaatioon samalla tavalla. Palkkasyrjintä on alhaisinta mediaanipalkan lähellä olevissa tuloluokissa ja suurinta hyvin matalasti tai korkeasti palkatuissa tehtävissä. Tutkielmassa saatiin myös epäsuoraa näyttöä siitä, että työnantajan vaihto muodostaa ylimääräisen riskin naisten palkkakehitykselle.

Palkkadiskriminaatio riippuu yli 20 tekijästä, joista sukupuolten erilainen ikäkohtelu on selvästi voimakkain yksittäinen syrjintää aiheuttava palkanmuodostuksen piirre. Miehet ja naiset aloittavat työuransa likimain samasta tilanteesta, mutta iän tuoma työkokemus kasvattaa miesten palkkaa huomattavasti nopeammin kuin naisten. Ero kumuloituu siten, että eläkeiän kynnyksellä diskriminaatio on melkein 14 log-%. Havaintoa tulkittaessa on huomioitava tutkielman laaja selittäjäjoukko. Mm. työn vaatavuutta ja tehtävää kontrolloitiin tarkasti. Niinpä

¹² Noin 8 tavallista prosenttia.

tutkielman saman työn käsite on sillä tavalla tiukka, että vertailu koskee yhtä pitkälle työurallaan edenneitä miehiä ja naisia – yleensähan ajatellaan, että uralla eteneminen tarkoittaa juuri siirtymistä haastavampiin ja korkeammin palkattuihin tehtäviin. Koska nämä työn piirteet on jo otettu huomioon, on havaittu miesten ja naisten erilainen ikäkohtelu luontevinta tulkita siten, että sukupuolia palkitaan *samasta* etenemisestä eri tavalla: miehillä se kanavoituu syystä tai toisesta suuremmaksi palkan nousuksi kuin naisilla.

On vaikeaa, ellei mahdotonta sanoa, mistä kahdeksan log-prosentin palkkadiiskriminaatio johtuu. On kehitetty lukuisia talousteoreettisia malleja, jotka osoittavat, että ainakin joissain tilanteissa palkkasyrjintä on työnantajan kannalta optimaalinen tapa toimia. Mm. Pasanen lähestyy palkkadiiskriminaatiota tästä näkökulmasta (1995). Teoreettiset mallit perustuvat kuitenkin aina joihinkin työnantajan ja -tekijän kannustimia ja mieltymyksiä koskeviin olettamuksiin. Koska niitä ei voida suoraan testata aineistosta, on jonkin teoreettisen selitysmallin paremmuus toiseen nähden pitkälti tulkintakysymys. Jokainen kannanotto vaatisi, että tehtäisiin (joskus voimakkaitakin) oletuksia työmarkkina-agenttien toiminnan perimmäisistä insentiiveistä. Niiden oikea spesifointi ei ole empiirinen, vaan erittäin haastava teoreettinen ongelma. Empiirinen tutkimus voi ainoastaan tarjota tietoa spesifioinnin tueksi ja tässä tutkielmassa on haluttu lähestyä palkkadiiskriminaatiota nimenomaan empiirisesti. Mm. Asplund tuo esiin teoreettisten selitysmallien moninaisuuden, niiden empiirisen testaamisen vaikeuden ja osittaisen provokatiivisuudenkin (1996, 5–10). Ei ole esimerkiksi mitenkään ongelmatonta väittää, että työnantajat maksaisivat jääkylmästi laskelmoiden naispuolisille työntekijöilleen miehiä alemmaa palkkaa. Yhtä pulmallinen on kanta, jonka mukaan työmarkkinoilla ei esiintyisi lainkaan tällaisia syrjiviä piirteitä – totuus lienee jossain kahden ääripään välissä. Lisäksi on huomioitava, että epäkokeelliseen aineistoon perustuva päättely on useimmiten luonteeltaan pois-sulkevaa. Hyvin tehdyin tilastollisin tarkasteluin voidaan kyllä osoittaa, että jokin seikka *ei* aiheuta palkkasyrjintää. Sen sijaan empirian keinoin on paljon vaikeampaa osoittaa pitävästi, että diskriminaatio todella aiheutuu jostain tekijästä – vaikkapa jonkin työmarkkina-agentin asenteista tai mieltymyksistä.

Edellä esitetyistä seikoista johtuen ei tässä haluta, eikä oikein voidakaan ottaa kantaa palkkasyrjinnän syihin. Eräs tutkielman tulos tarjoaa kuitenkin mielenkiintoisen yhtymäkohdan keskusteluun, jota julkisuudessa on käyty palkkasyrjinnän taustalta löytyvistä mekanismeista. Varsin yleisesti nimittäin arvellaan, että palkkadiiskriminaatio johtuisi naisten äitiyslomien vuoksi katkonaisesta työurasta. Tämän näkökulman mukaan työnantajat toimisivat rationaalisesti ja maksaisivat naisille samasta työstä miehiä alemmaa palkkaa, koska heihin liittyy miehiä suurempi riski pitkistä poissaoloista. (Asplund 1996, 10.) Kohdassa 5.6 havaittiin, että naiset tarvitsisivat noin kuusi vuotta enemmän työkokemusta, jotta he muuten samanlaisessa työssä saisivat samaa palkkaa kuin miehet. Lukema on hämmästyttävän lähellä aikaa, jonka voidaan luontevasti arvioida kuluvan kahden pienen lapsen hoitoon – ja Suomessa hankitaan keskimäärin juuri kaksi lasta. On

painotettava, että osoitettu yhteys on puhtaasti laskennallinen tulos ja siis vailla kausaalitulkintaa. Tilastollisin tarkasteluin ei voida saada selville, miksi naiset ovat jääneet yli kuudella työkokemusvuodella jälkeen miesten palkkakehityksestä. Silti havaittu yhteys on merkille panemisen arvoinen. Tulevaisuudessa olisikin kiinnostavaa tutkia, mitkä ovat todelliset työnantajalle äitiyslomista aiheutuvat kustannukset. Jos osoittautuu, että palkkadiskriminaation suuruus voidaan suurimmaksi osaksi palauttaa niihin, korostaisi tulos tarvetta jakaa lasten hankinnasta aiheutuvat kustannukset nykyistä tasaisemmin sekä äidin että isän työnantajalle.

Toisaalta jatkotarkasteluissa saattaa osoittautua, että lastenhoidon vuoksi naisilla todella on vähemmän työkokemusta kuin saman ikäisillä miehillä. Tällöin lasten hankinnasta koituisi naisille ylimääräisiä kustannuksia menetettynä työkokemuksena ja palkkadiskriminaatio heijastelisi tämän inhimillisessä pääomassa mitattavan kustannuksen erittäin epätasaista jakautumista sukupuolten kesken. Ongelmallista on, että näitä kustannuksia on hankala havaita, koska niitä on vaikea mitata rahassa. Ehkä siksi niihin myös kiinnitetään vähemmän huomiota kuin välittömiin työnantajalle aiheutuviin kuluihin. Lisäksi ne realisoituvat viipeellä: koska juuri pätevyys omassa toimessa on keskeinen ylennyksiä ja palkankorotuksia tuova työntekijän ominaisuus, kasvaa menetety työkokemuksen aiheuttama palkkatappio kumulatiivisesti työuran edetessä¹³. Jos vastaisuudessa saadaan tällaista hypoteesia tukevia empiirisiä tuloksia, korostaisivat ne tarvetta jakaa myös tämä, pitkälti epäsuora lasten hankkimisesta aiheutuva kustannus nykyistä tasaisemmin sukupuolten kesken. Tällöin keskeiseen asemaan nousisi miesten runsaampi osallistuminen lasten hoitoon vaikkapa joustavampien työaikojen ja osapäivätyön kautta. Toisaalta pienten lasten äitien osallistumista työelämään tulisi samoin keinoin edistää. Perimmältään palkkadiskriminaatio palautuisi siis siihen, kuinka hoiva yhteiskunnassamme organisoidaan.

Lopuksi on todettava, että edellä esitetyt (potentiaaliset) selitysvaihtoehdot eivät sulje toisiaan pois, vaan molemmat voivat pitää samanaikaisesti paikkansa. Ne edellyttävät kuitenkin kovin erilaisia toimenpiteitä. Siksi palkkadiskriminaation pienentämiseen parhaiten soveltuvat keinot riippuvat ratkaisevasti siitä, kumpi mekanismeista on voimakkaampi. Jatkossa olisikin erityisen tärkeää selvittää niiden suhteellinen merkitys, jos palkkasyrjinnän taustalta todella löytyy lasten hankinnasta aiheutuvien kustannusten epätasainen jakautuminen.

¹³ Jos naisilla on miehiä katkonaisempi työura, tukisi kuvio 10 hypoteesia menetety työkokemuksen aiheuttamasta kumulatiivisesta tappiosta.

Lähteet

- Asplund, R.* (1993a): Essays on Human Capital and Earnings in Finland. Elinkeinoelämän Tutkimuslaitos A18. Taloustieto, Helsinki.
- Asplund, R.* (1993b): Teollisuuden toimihenkilöiden palkat ja inhimillinen pääoma. Elinkeinoelämän Tutkimuslaitos B89. Taloustieto, Helsinki.
- Asplund, R.* (1996): Koulutus, työura ja palkkaerot. Elinkeinoelämän Tutkimuslaitoksen keskusteluaiheita No. 553. Taloustieto, Helsinki.
- Davidson, R. – Mackinnon, J. G.* (1993): Estimation and inference in econometrics. Oxford University Press, New York.
- Dougherty, C.* (1992): Introduction to Econometrics. Oxford University Press, New York.
- Ekholm, A.* (2000): Johdatus uskottavuuspäätelyyn. 2. painos. Helsingin yliopiston tilastotieteen laitoksen opetusmoniste 1/2000.
- Finlex* (2003): Valtion Säädöstietopankki Finlex. www.finlex.fi. Viitattu 15.1, 16.1, 21.1, 22.1, 23.1, 28.1 ja 7.2.2003.
- Groshen, E. L.* (1991): The Structure of the Female/Male Wage Differential. The journal of Human resources, Vol 26, No. 3, 457–472. The University of Wisconsin Press.
- Hauhio, N. – Lilja, R.* (1996): The Evolution of Gender Wage Differentials over the Career. Elinkeinoelämän Tutkimuslaitoksen keskusteluaiheita No. 573. Taloustieto, Helsinki.
- Kettunen, J. – Marjanen, R.* (1992): Suomen Työnantajain Keskusliiton palkkatilastot: Syntyhistoria, sisältö ja käyttötarkoitus. Elinkeinoelämän Tutkimuslaitoksen keskusteluaiheita No. 425. Helsinki.
- Kloek, T.* (1981): OLS Estimation in a Model where a Microvariable Is Explained by Aggregates and Contemporaneous Disturbances Are Equicorrelated. *Econometrica*, Vol 49, No. 1, 205–207.
- Koiv, E.* (1996): Palkkadiskriminaatio teollisuuden toimihenkilöillä. Elinkeinoelämän tutkimuslaitoksen keskusteluaiheita No. 568. Taloustieto, Helsinki.
- Korkeamäki, O. – Kyyrä, T.* (2002): The Gender Wage Gap and Sex Segregation in Finland. Valtion taloudellisen tutkimuskeskuksen keskustelualoitteita nro 288. Helsinki.
- Lilja, R.* (1995): Teollisuuden toimihenkilöiden ura- ja palkkakehitys. Palkansaajien tutkimuslaitoksen tutkimuksia nro 54. Hakapaino, Helsinki.

- Lilja, R.* (1996): Toimihenkilöura Suomen teollisuudessa. Elinkeinoelämän Tutkimuslaitos B117. Taloustieto, Helsinki.
- Lilja, R.* (1997): Similar Education – Different Career and Wages? Elinkeinoelämän Tutkimuslaitoksen keskusteluaiheita No. 606. Taloustieto, Helsinki.
- Moulton, B. R.* (1990): An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units. *The Review of Economics and Statistics*, Vol 72, No. 2, 334–338.
- Oaxaca, R.* (1973): Male – Female Wage Differentials in Urban Labour Markets. *International Economic Review*, Vol 14, No. 3, 693–709.
- Pasanen, J.* (1995): Syrjinnän talousteoria ja naisten palkkasyrjinnän suuruus Suomen metsä- ja metalliteollisuuden suuryrityksissä. Pro gradu-tutkielma. Helsingin yliopiston valtiotieteellinen tiedekunta.
- Pohjanpalo, O.* (2003a): Palkkatasa-arvoa ei vaadita suureen ääneen. Artikkel. Helsingin sanomat 10.3.2003, s. A6 / politiikka.
- Pohjanpalo, O.* (2003b): Naisehdokkaat hakivat puhtia vaalikampanjaansa. Artikkel. Helsingin sanomat 10.3.2003, s. A6 / politiikka.
- Spanos, A.* (1986): *Statistical Foundations of Econometric Modelling*. Cambridge University Press, Cambridge.
- TT* (2001): Palkkatilastokatsaus 2000. Teollisuuden ja Työnantajain Keskusliitto, työmarkkinatilastot 2001. Helsinki.
- TT* (2003a): Teollisuuden ja Työnantajain Keskusliiton internet-sivut. www.tt.fi. Viitattu 3.12, 10.12.2002, 2.1 ja 3.1.2003.
- TT* (2003b): Naiset ja miehet työmarkkinoilla. Teollisuuden ja Työnantajain Keskusliiton tiedote 10.4.2003. www.tt.fi/arkisto/getoriginal.pl?ft_cid=3795. Viitattu 12.4.2003.
- Vartia, Y.* (1988): Ovatko tilastolliset tutkimustulokset vain monimutkaisesti perusteltuja mielipiteitä? *Kansantaloudellinen aikakauskirja* 1988:3, 305–309.
- Vartia, Y.* (1995): Promootiokirjan esitelmä ja promoottorin puhe: Palkkadiskriminaatio. Helsingin yliopiston kansantaloustieteen laitoksen esitelmät ja tutkimuksia 11/1995.
- Vartia, Y. – Koskimäki, T.* (2001): Beyond Matched Pairs and Griliches-Type Hedonic Methods for Controlling Quality Changes in CPI Sub-indices. Helsingin yliopiston kansantaloustieteen laitoksen keskustelualoitteita No. 523.
- Vartia, Y. – Kurjenoja, J.* (1992): Palkkadiskriminaatio. Naisten ja miesten palkkaerot samasta työstä metalli- ja metsäteollisuuden suuryrityksissä v.

1990. Helsingin yliopiston kansantaloustieteen laitoksen tutkimuksia, No. 60:1992.

Vartiainen, J. (1998): Naisten ja miesten palkkaero Suomen metalliteollisuudessa. Palkansaajien tutkimuslaitoksen tutkimuksia nro 71. Hakapaino, Helsinki.

Vartiainen, J. (1999): Job Assignment and the Gender Wage Differential: Theory and Evidence on Finnish Metalworkers. Palkansaajien tutkimuslaitoksen tutkimusselosteita nro 155. Hakapaino, Helsinki.

Vartiainen, J. (2001): Sukupuolten palkkaeron tilastointi ja analyysi. Sosiaali- ja terveysministeriön tasa-arvojulkaisuja 2001:7. Edita, Helsinki.

White, H. (1980): A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, Vol 48, No. 4, 817–838.

Wonnacott, R. J. – Wonnacott, T. H. (1979): *Econometrics*. John Wiley & Sons, New York.

LIITE 1. Sopimusalat ja niillä käytettävät työn vaativuusasteikot

<i>Sopimusala</i>	<i>Osuus ai- neistosta, %</i>	<i>Asteikkoja, kpl</i>	<i>Portaita alalla käytetyissä asteikoissa, kpl</i>
Metalli- ja elektroniikkateollisuus	34,56	3	9, 10 tai 15
Posti	10,40	1	8
Paperiteollisuus	5,59	2	6 tai 12
Muovi- ja kemian tuoteteollisuus	4,39	1	8
Teleliikenneala	4,29	1	9
Tietoliikenneala	4,20	1	9
Graafinen teollisuus	4,01	2	6 tai 12
Suunnittelu- ja konsulttiala	3,78	2	6 tai 12
Talonrakennusala	3,26	2	6 tai 12
Meijerit	3,05	1	14
Energia-ala	2,93	2	6 tai 12
Leipomot	2,44	1	14
Kemian perusteollisuus	2,13	1	8
Saha- ja levyteollisuus	1,64	2	6 tai 12
Peruselintarviketeollisuus	1,53	1	14
Rakennustuoteteollisuus	1,39	1	9
Öljy- kaasu- ja petrokemian teollisuus	1,20	1	8
Metsätyönjohtajat	1,19	1	3
Lihanjalostusteollisuus	1,06	1	14
Tekstiiliteollisuus	0,68	1	10
Panimoteollisuus	0,66	1	14
Lasikeraaminen teollisuus	0,64	1	10
Puusepänteollisuus	0,62	2	6 tai 12
Sähköasennusala	0,61	2	6 tai 12
Metsäosastot	0,57	2	6 tai 12
Öljytuote- ja säiliöautoala	0,52	1	7
Vaatetusteollisuus	0,46	1	10
Ahtausala	0,37	2	6 tai 12
Putkiasennusala	0,36	2	6 tai 12
Kumiteollisuus	0,34	1	10
Vedeneristys- ja tienpäällystysala	0,28	2	6 tai 12
Pesula-ala	0,22	1	8
Maa- ja vesirakennusala	0,17	2	6 tai 12
Kiinteistöhoitoala	0,15	2	6 tai 12
Pahvi- ja paperituoteteollisuus	0,15	2	6 tai 12
Kenkäteollisuus	0,09	1	9
Nahkateollisuus	0,03	1	10
Rakennuspelti- ja teollisuuseristysala	0,03	1	9

**VATT-KESKUSTELUALOITTEITA / DISCUSSION PAPERS ISSN 0788-5016
- SARJASSA ILMESTYNEITÄ**

252. Bjerstedt Katja: Työssä jaksamisesta ja työmarkkinoiden muutoksesta. Helsinki 2001.
253. Sinko Pekka: Unemployment Insurance with Limited Duration and Variable Replacement Ratio – Effects on Optimal Search. Helsinki 2001.
254. Rauhanen Timo: Arvonlisäverotus EU:n jäsenmaissa – voiko vientiä verottaa? Helsinki 2001.
255. Korkeamäki Ossi: Työttömyysriskiin vaikuttavat yksilö- ja yrityskohtaiset tekijät Suomessa 1990-1996. Helsinki 2001.
256. Kyyrä Tomi: Estimating Equilibrium Search Models from Finnish Data. Helsinki 2001.
257. Moisio Antti: On Local Government Spending and Taxation Behaviour – effect of population size and economic condition. Helsinki 2001.
258. Kari Seppo – Ylä-Liedenpohja Jouko: Klassillinen osakeyhtiövero kansainvälisen veroharmonisoinnin muotona. Helsinki 2001.
259. Kiander Jaakko – Vaittinen Risto: EU:n itälaajenemisen vaikutuksista: laskelmia tasapainomallilla. Helsinki 2001.
260. Hakola Tuulia – Uusitalo Roope: Let's Make a Deal – the Impact of Social Security Provisions and Firm Liabilities on Early Retirement. Helsinki 2001.
261. Hjerppe Reino – Kari Seppo – Lönnqvist Henrik (toim.): Verokilpailu ja verotuksen kehittämistarpeet. Helsinki 2001.
262. Hakola Tuulia – Lindeboom Maarten: Retirement Round-about: Early Exit Channels and Disability Applications. Helsinki 2001.
263. Kilponen Juha – Romppanen Antti: Julkinen talous ja väestön ikääntyminen pitkällä aikavälillä – katsaus kirjallisuuteen ja simulointeja sukupolvimallilla. Helsinki 2001.
264. Riihelä Marja – Sullström Risto – Tuomala Matti: On Economic Poverty in Finland in the 1990s. Helsinki 2001.
265. Parkkinen Pekka: Suomen ja muiden unionimaiden väestön ikärakenne vuoteen 2050. Helsinki 2002.
266. Kari Seppo – Ylä-Liedenpohja Jouko: Classical Corporation Tax as a Global Means of Tax Harmonization. Helsinki 2002.
267. Kallioinen Johanna: Pyöräilyn institutionaalinen asema liikennesuunnittelussa. Helsinki 2002.
268. Kangasharju Aki and Venetoklis Takis: Business Subsidies and Employment of Firms: Overall Evaluation and Regional Extensions. Helsinki 2002.
269. Moisio Antti: Determinants of Expenditure Variation in Finnish Municipalities. Helsinki 2002.
270. Riihelä Marja – Sullström Risto: Käytettävissä olevien tulojen liikkuvuus vuosina 1990-1999. Helsinki 2002.

271. Kari Seppo – Kröger Outi – With Hanna: Saksan verouudistus 2001. Helsinki 2002.
272. Kari Seppo: Norjan ja Suomen tuloverojärjestelmän vertailua. Helsinki 2002.
273. Sinko Pekka: Labour Tax Reforms and Labour Demand in Finland 1997-2001. Helsinki 2002.
274. Pekkala Sari – Kangasharju Aki: Adjustment to Regional Labour Market Shocks. Helsinki 2002.
275. Poutvaara Panu: Gerontocracy Revisited: Unilateral Transfer to the Young May Benefit the Middle-aged. Helsinki 2002.
276. Uusitalo Roope: Tulospalkkaus ja tuottavuus. Helsinki 2002.
277. Kemppi Heikki – Lehtilä Antti: Hiilidioksidiveron taloudelliset vaikutukset. Helsinki 2002.
278. Rauhanen Timo: Liikevaihtoraja arvonnäköverotuksessa – Pienyrityksen kasvun este? Helsinki 2002.
279. Berghäll Elina – Kilponen Juha – Santavirta Torsten: Näkökulmia suomalaiseen tiede-, teknologia- ja innovaatiopolitiikkaan – KNOGG-työpajaseminaarin yhteenveto. Helsinki 2002.
280. Laine Veli: Evaluating Tax and Benefit Reforms in 1996 – 2001. Helsinki 2002.
281. Pekkala Sari – Tervo Hannu: Unemployment and Migration: Does Moving Help? Helsinki 2002.
282. Honkatukia Juha – Joutsenvirta Esa – Kemppi Heikki – Perrels Adriaan: EU:n laajuuden päästökaupan toteuttamisvaihtoehdot ja vaikutukset Suomen kannalta. Helsinki 2002.
283. Kotakorpi Kaisa: Access Pricing and Competition in Telecommunications. Helsinki 2002.
284. Hämäläinen Kari – Böckerman Petri: Regional Labour Market Dynamics, Housing and Migration. Helsinki 2002.
285. Sinko Pekka: Labour Taxation, Tax Progression and Job Matching – Comparing Alternative Models of Wage Setting. Helsinki 2002.
286. Tuomala Juha: Työttömyyden alueellisen rakenteen kehitys 1990-luvulla. Helsinki 2002.
287. Berghäll Elina: Virosta ja Ahvenanmaan itsehallintoalueelta Suomeen kohdistuvat verokilpailupaineet. Helsinki 2002.
288. Korkeamäki Ossi – Kyyrä Tomi: The Gender Wage Gap and Sex Segregation in Finland. Helsinki 2002.
289. Kilponen Juha – Santavirta Torsten: Taloudellinen kasvu, innovaatiot ja kilpailu – katsaus kirjallisuuteen ja politiikkasuosituksiin. Helsinki 2002.
290. Siivonen Erkki – Huikuri Satu (Edited): Workshop on Studies for Northern Dimension Kalastajatorppa 30 – 31 May, 2002. Helsinki 2002.
291. Pohjola Johanna – Kerkelä Leena – Mäkipää Raisa: Who Gains from Credited Forest Carbon Sinks: Finland and Other Annex I Countries in Comparison. Helsinki 2002.

292. Montén Seppo – Tuomala Juha: Alueellinen työttömyys ja pitkäaikaistyöttömyys 1990-luvulla. Helsinki 2003.
293. Lyytikäinen Teemu: Pienituloisuuden dynamiikka Suomessa. Helsinki 2003.
294. Aulin-Ahmavaara Pirkko – Jalava Jukka: Pääomapanos ja sen tuottavuus Suomessa vuosina 1975-2001. Helsinki 2003.
295. Vaittinen Risto: Maatalouskaupan vapauttaminen – kansainväliset vaikutukset ja merkitys EU:lle. Helsinki 2003.
296. Haataja Anita: Suomalaiset mikrosimulointimallit päätöksenteon valmistelussa ja tutkimuksessa. Helsinki 2003.
297. Kangasharju Aki – Korpinen Liisa – Parkkinen Pekka: Suomessa asuvat ulkomaalaiset: Esiselvitys. Helsinki 2003.
298. Hietala Harri – Lyytikäinen Teemu: Työn, pääoman ja kulutuksen verorasituksen mittaaminen. Helsinki 2003.
299. Räisänen Heikki: Rekrytointiongelmien ja työvoimapotentialin lääkärien, lastentarhanopettajien, farmaseuttien ja proviisorien ammateissa. Helsinki 2003.
300. Kröger Outi: Pääoma- ja yritystulojen verotus – uusi suunta? Helsinki 2003.
301. Kari Seppo – Liljeblom Eva – Ylä-Liedenpohja Jouko: Snedvridande beskattning av utländska investeringar: Reell och finansiell aktivitet inducerad av skattearbitrage. Helsinki 2003.
302. Pekkala Sari: Is Little Brother Nothing but Trouble?: Educational Attainment, Returns to Schooling and Sibling Structure. Helsinki 2003.
303. Vaittinen Risto: Liberalisation of Agricultural Trade – Global Implications and what it Means for the EU. Helsinki 2003.
304. Kangasharju Aki – Venetoklis Takis: Do Wage-subsidies Increase Employment in Firms? Helsinki 2003.
305. Räisänen Heikki: How to Cope with Labour Market Policy Tools in Economic Downturn: Finnish Evidence. Helsinki 2003.
306. Ruotoistenmäki Riikka – Siivonen Erkki: Tiehankkeiden rahoitusvajeen ratkaisu? Helsinki 2003.
307. Hjerpe Reino: Social Capital and Economic Growth Revisited. Helsinki 2003.
308. Honkatukia Juha – Kangasharju Aki – Vaittinen Risto: Suuren aluepolitiikan ja hajasijoittamisen vaikutuksia Keski-Suomessa. Helsinki 2003.