VATT-TUTKIMUKSIA
130
VATT RESEARCH REPORTS

Jussi Lintunen

TULOEROJEN JA TALOUDELLISEN ERIARVOISUUDEN MITTAAMISESTA: SOVELLUS SUOMEN KULUTUSTUTKIMUKSILLA

Valtion taloudellinen tutkimuskeskus
Government Institute for Economic Research
Helsinki 2007

ISBN 978-951-561-701-9 (nid.)
ISBN 978-951-561-702-6 (PDF)

ISSN 0788-5008 (nid.)
ISSN 1795-3340 (PDF)

Valtion taloudellinen tutkimuskeskus
Government Institute for Economic Research
Arkadiankatu 7, 00100 Helsinki, Finland

Email: jussi.lintunen@helsinki.fi

Oy Nord Print Ab
Helsinki, maaliskuu 2007

JUSSI LINTUNEN: TULOEROJEN JA TALOUDELLISEN ERIARVOISUUDEN MITTAAMISESTA: SOVELLUS SUOMEN KULUTUSTUTKIMUKSILLA. Helsinki, VATT, Valtion taloudellinen tutkimuskeskus, Government Institute for Economic Research, 2007, (B, ISSN 0788-5008 (nid.), ISSN 1795-3340 (PDF), No 130). ISBN 978-951-561-701-9 (nid.), ISBN 978-951-561-702-6 (PDF).

Tiivistelmä: Tutkimuksessa tarkastellaan kolmen yleisesti käytetyn eriarvoisuusmitan, laajennetun Gini-kertoimen, Atkinsonin indeksin sekä yleistetyn entropian, ominaisuuksia tuloerojen kuvaajina. Tavoitteena on havainnollistaa mittojen erilaisia painotuksia jakauman eri osille. Tämä mahdollistaa eriarvoisuustulkinnan antamisen tuloksille. Empiirisessä osiossa tutkitaan tuloerojen kehitystä Suomessa vuosien 1985 – 2001 välisenä aikana Tilastokeskuksen Kulutustutkimuksia käyttäen. Tarkastelut tehdään käyttäen laajennettua Gini-kerrointa painottaen jakauman osia eri tavoin. Painotettaessa suurituloisia havaitaan tuotannontekijätuloissa voimakasta tuloerojen kasvua 1990-luvun alkupuoliskolla, joka on sittemmin tasaantunut. Käytettävissä olevissa tuloissa tämä näkyy tasaisena, mutta hidastuvana, tuloerojen kasvuna. Painotettaessa pienituloisia, havaitaan tuotannontekijätuloissa voimakasta tuloerojen kasvua 1990-luvun alkupuoliskolla, jonka jälkeen tuloerot ovat olleet laskussa. Käytettävissä olevissa tuloissa tuloerojen kasvu näkyy kuitenkin vasta 1990-luvun loppupuoliskolla.

Asiasanat: tuloerot, eriarvoisuus, hyvinvointi, mittaus

JUSSI LINTUNEN: TULOEROJEN JA TALOUDELLISEN ERIARVOISUUDEN MITTAAMISESTA: SOVELLUS SUOMEN KULUTUSTUTKIMUKSILLA. Helsinki, VATT, Valtion taloudellinen tutkimuskeskus, Government Institute for Economic Research, 2007, (B, ISSN 0788-5008 (nid.), ISSN 1795-3340 (PDF), No 130). ISBN 978-951-561-701-9 (nid.), ISBN 978-951-561-702-6 (PDF).

Abstract: In this study three commonly used inequality indices, extended Gini coefficient, Atkinson index and generalized entropy index, are examined. The aim is to illustrate the different weights the indices give to different parts of the distribution. This makes it possible to make normative valuations. In the empirical part income inequality in Finland in a period of 1985 – 2001 is studied using Household Budget Survey by Statistics Finland. The measure used here is the extended Gini coefficient. Giving more weight to the high income end of factor income distribution, there is strong growth in income inequality in the early 1990s, but it has leveled since. In disposable income, growth has been steady but not as strong. Giving weight to low income part, there was a very strong increase in income inequality in the early 1990s but is has lowered after 1995. In the case of disposable income, the increase is delayed until late 1990s.

Key words: income inequality, inequality, welfare, measurement

Saatteeksi

Tutkimus perustuu Helsingin yliopistossa lokakuussa 2006 tarkastettuun ja hyväksyttyyn pro gradu-tutkielmaan. Työn ohjaajana toimi professori Yrjö Vartia. Kiitän häntä aiheen tiimoilta käymistämme hyödyllisistä keskusteluista.

Kiitän myös Valtion taloudellista tutkimuskeskusta harjoittelupaikasta kesällä 2005 ja mahdollisuudesta tehdä tutkimusta sen tiloissa lukuvuonna 2005 – 2006. Erityiskiitos tutkimusjohtaja Heikki Räisäselle edellä mainitun mahdollistamisesta. Lisäksi tahdon kiittää erikoistutkija Marja Riihelää ja erikoistutkija Risto Sullströmiä antoisista keskusteluista, ohjelmointiavusta sekä tulonjakokysymysten ääreen johdattamisesta.

Työn mahdollisista virheistä ja puutteista vastaan luonnollisesti itse.

Helsingissä 2. maaliskuuta 2007

Jussi Lintunen

Sisällys

1 Johdanto

1.1 Tuloerotutkimuksen lähtökohdista

Tutkittaessa tuloeroja ja taloudellista eriarvoisuutta lähtökohtana on havaittu tulojakauma. Koska tulon ajatellaan tuottavan saajalleen taloudellista hyvinvointia, voidaan tulojakauman ajatella generoivan jonkinlaisen taloudellisen hyvinvoinnin jakauman. Samalla tulojakaumassa havaitut tuloerot generoivat hyvinvointijakaumaan hyvinvointierot. Tässä työssä tarkastellaan näitä tulo- ja hyvinvointieroja. Jotta tuloerojen ja eriarvoisuuden mittaamisen teorian voisi ymmärtää, on tutustuttava ongelman peruskäsitteisiin.

1.1.1 Tulot ja taloudellinen hyvinvointi

Kotitaloudet ansaitsevat erilaisia tuloja. Kansantalouden tilinpitojärjestelmässä nämä tulot on luokiteltu palkka-, yrittäjä- sekä omaisuustuloiksi. Näiden tuotannontekijätulojen lisäksi kotitaloudet *saavat* etenkin julkiselta sektorilta sekä veronalaisia että verottomia tulonsiirtoja. Näin saaduista tuloista (bruttotuloista) kotitaloudet *maksavat* veroja sekä veroluonteisia maksuja, so. maksettuja tulonsiirtoja. Jäljelle jäävää osuutta kutsutaan käytettävissä oleviksi tuloiksi (nettotuloksi), ja sen kotitalous voi käyttää kuluttamiseen tai säästämiseen. Lisäksi voidaan kuluttaa yhteiskunnallisia palveluita, jotka usein puuttuvat tilastoaineistoista.

Taloudellisen hyvinvoinnin ajatellaan koostuvan kulutuksen tuottamista hyödyke- ja palveluvirroista. Kulutus on mahdollista, jos saadut tulot tai aiemmin kertyneet säästöt sen sallivat.[1] Siten käytettävissä olevat tulot mahdollistavat niin nykyisen kuin tulevankin kulutuksen. Käytettävissä olevat tulot liittyvät siis läheisesti kotitalouden kokemaan taloudelliseen hyvinvointiin.

Tässä tutkimuksessa oletetaan, että käytettävissä olevat tulot määrittävät yksikäsitteisesti kotitalouden taloudellisen hyvinvoinnin tason. Kokonaishyvinvointi on tietenkin useamman tekijän summa, ja siksi nyt tehdyt tarkastelut kohdistuvat vain yhteen hyvinvoinnin osa-alueeseen.

Teoriassa yksilön taloudellisen hyvinvoinnin tason määrää hänen elinikänsä aikana ansaitsemansa tulot eli elinkaaritulot. Nämä tulot määrittävät yksikäsitteisesti henkilön kulutusmenojen budjettirajoitteen. Kotitalouksia tutkittaessa on huomioitava kotitalouden kokonaisuutena ansaitsemat elinkaaritulot. Tämä ei käytännössä ole mahdollista, sillä elinkaarituloista ei ole havaintoaineistoja. Tämän vuoksi tutkimus tehdään kotitalouden yhden vuoden tulojen perusteella.

[1] Lainarahoitus on toki mahdollista, mutta luottokelpoisuuden määrää menneet ja tulevat tulot.

1.1.2 Tasa-arvo ja eriarvoisuus

Tasa-arvon periaatteiden mukaan samanlaisia ihmisiä on kohdeltava samalla tavalla. Siten samalla tavalla samaa työtä samassa yrityksessä tekevien henkilöiden pitäisi saada samaa palkkaa. Ei ole siis tasa-arvoista maksaa samanlaisille eri määrää tuloa. Mutta ei ole myöskään tasa-arvoista maksaa erilaisille samaa määrää tuloa. Ei liene epäilystä siitä, että opiskelijatalouden tulojen on oltava pienemmät kuin suuryrityksen johtajan. Tärkeä kysymys on siis se, kuinka suuret tuloerot ovat oikeudenmukaisia? Milloin erilaisuudesta saadut lisätulot tai menetetyt tulot ovat niin suuria, että tulonjakoa ei enää voi pitää tasa-arvoisena. Kysymys on tärkeä, mutta myös vaikea.

Jotta edellä esitettyä kysymystä voitaisiin alkaa pohtia, on ensin mietittävä, mitä tarkoitetaan samanlaisuudella? Mitkä ominaisuudet ovat sellaisia, että niiden eroista on oikeutettua saada korvaus? Yleisesti hyväksyttyä on, että sukupuoli ja ihonväri ovat ominaisuuksia, jotka eivät saisi vaikuttaa tulojen määrään. Toisaalta koulutus, ahkeruus ja työtehtävä näyttävät oikeuttavan erisuuruisiin tuloihin. Kun samanlaisuus ja siten erilaisuus on saatu määriteltyä, voidaan pohtia sitä, mikä on vallitseva sosiaalinen normi, joka määrää oikeudenmukaisen korvauksen erilaisuudelle. Näihin kysymyksiin ei tässä työssä anneta vastausta.

1.1.3 Tuloerot ja taloudellinen eriarvoisuus

Kotitaloudet saavat eri määrän tuloja. Tyypillisesti tulot ovat positiivisia ja jakaumalla on pitkä häntä suurilla tulojen arvoilla. Tämä tulojakauma generoi hyötyfunktion määrittelemän hyötyjakauman. Tulojakaumavertailut perustuvat tähän hyötyjakaumaan, koska se pitää sisällään informaation yhteiskunnan taloudellisesta hyvinvoinnista, jonka perusteella jakauman hyvyyttä arvioidaan. Jakaumavertailujen teoriaa esitellään luvussa 2.

Tuloerot aiheuttavat sen, että eri kotitalouksilla on erilainen taloudellinen hyvinvointi. Jatkossa tehdyt tarkastelut pyrkivät mittaamaan tätä hyvinvoinnin hajontaa, ja sitä kutsutaan jatkossa (taloudelliseksi) eriarvoisuudeksi. Näin työssä tutkittu eriarvoisuus ei vastaa täysin tasa-arvokäsityksen mukaista eriarvoisuutta, vaan osa havaituista tuloeroista on oikeutettuja ja osa mahdollisesti ei. Ongelmallisen eriarvoisuuskäsitteen käyttöön on kaksi syytä: Emme tiedä kotitalouksien ominaisuuksia riittävällä tarkkuudella, jotta tietäisimme oikeutuksen tuloeroille. Emme myöskään tiedä, mikä on sosiaalinen normi oikeutetuille tuloeroille. Siten tutkitun eriarvoisuuden muutokset voivat johtua siitä, että kotitalouksien ominaisuuksien jakauma on muuttunut tai että edellä mainittu normi on muuttunut. Lisäksi havainnoissa on tietenkin mukana myös todellisen eriarvoisuuden muutos.

Lisäksi on syytä painottaa sitä seikkaa, että ei ole perusteltua pitää optimaalisena havaitun eriarvoisuuden häviämistä. Tämä tarkoittaisi juuri sitä, että erilaisille maksettaisiin saman verran tuloa, mikä ei ole oikeudenmukaista. Tämän vuoksi optimaalisen havaitun taloudellisen eriarvoisuuden määrä poikkeaa nollasta.

1.2 Eriarvoisuusmitoista

Tuloeroja ja eriarvoisuutta voidaan mitata ns. eriarvoisuusmitoilla. Näitä mittoja on määritelty lukuisia ja niiden perusteluiksi on esitetty erilaisia aksiomatisointeja. Täten ei liene yllätys, että eri mitat voivat järjestää jakaumat eri tavalla. Sanotaan, että ne eivät ole keskenään järjestysekvivalentteja. Käytetyn eriarvoisuusmitan valinta sisältää siis *arvovalinnan*, josta on syytä olla tietoinen.

Mitat on tyypillisesti suunniteltu teoreettisille tulojakaumille, jotka vastaavat reaalimaailmassa lähinnä elinkaarituloja. Tehtäessä tutkimusta vuosituloilla on syytä huomata, että tulojen vuosittaisvaihtelu voi olla voimakasta. Tästä johtuen kaikki havaittu tulojen hajonta ei välttämättä johda hyvinvoinnin hajontaan. Lisäksi käytetyissä kulutustutkimusten aineistoissa on vuosittaisvaihtelusta johtuvia sekä negatiivisia että nollahavaintoja. Tämä on ongelma monille eriarvoisuusmitoille.

Edelleen, mitat on tyypillisesti normitettu siten, että tasaisesti jaettu tulo antaa mitalle arvon nolla, jolloin eriarvoisuutta ei mitan perusteella tällöin ole. Kuten edellä mainittiin, ei tällainen jakauma olisi tasa-arvoinen. Voidaankin sanoa, että mittojen optimaalinen arvo poikkeaa nollasta, mutta sen tarkkaa arvoa on vaikea määrittää. Mittojen optimaaliseen absoluuttiseen arvoon liittyvästä ongelmasta huolimatta voidaan mittojen arvoissa tapahtuneita ajallisia muutoksia tutkia. Muutosten lähteet ovat tuotannontekijämarkkinoilla tapahtuneet muutokset sekä julkisen vallan tulonjakotoimissa tapahtuneet muutokset. Kun eri mittojen arvojen muutokset on saatu selvitettyä, voi *mittojen arvovalinnat tunteva* tarkastelija antaa muutoksille eriarvoisuustulkinnan. Mitat kuvaavat siis tuloerojen suuruutta, ja eriarvoisuustulkinta on tähän perustuva normatiivinen kysymys.

1.3 Motivaatio ja kysymyksen asettelu

Gini-kertoimella mitattuna tuloerot ovat Suomessa merkittävästi kasvaneet vuosina 1985–2001.[2] On mielenkiintoista tutkia, miltä tämä muutos näyttää käytettäessä eri eriarvoisuusmittaperheitä erilaisilla parametriarvoillaan. Tutkittaessa lisäksi mittojen arvoja eri tulokäsitteillä voidaan havaita, missä määrin muutokset ovat aiheutuneet muutoksista tuotannontekijämarkkinoilla ja missä määrin julkisen vallan tulontasaustoimien vaikutuksen muutoksista.

Jotta edellä kuvattu tutkimus voidaan tehdä ja tuloksia tulkita, on selvitettävä, millaisia arvovalintoja mittoihin sisältyy. Tutkimuksessa perehdytään kolmeen eniten käytettyyn mittaperheeseen: laajennettuun Gini-kertoimeen, Atkinsonin indeksiin sekä yleistettyyn entropiaan. Työssä tarkastellaan sitä, missä määrin mitat pitävät eriarvoisuutta lisäävänä pienituloisten ja missä määrin suurituloisten suurta osuutta. Lisäksi havainnollistetaan mittojen parametriarvojen vaikutusta kyseisiin painotuksiin. Kun edellä mainitut seikat ovat selvillä, voidaan pohtia, mikä mitta ja millä parametrin arvolla vastaa parhaiten käsitystämme reilusta painotuksesta. Näin eriarvoisuusmittojen arvoille ja arvojen muutoksille voi antaa mielekkään eriarvoisuustulkinnan.

[2] Gini-kerroin on kasvanut yli 5 prosenttiyksikköä, joka on noin 25 % lähtötasosta (esim. Riihelä et al. 2005).

2 Jakaumavertailut

Jakaumia voidaan vertailla ja järjestää monien ominaisuuksien, kuten esimerkiksi muodon, sijainnin tai hajonnan, perusteella. Eri ominaisuudet johtavat erilaisiin menetelmiin ja tunnuslukuihin. Tulojakaumia verrattaessa pyritään usein asettamaan jakaumat yhteiskunnallisen paremmuuden mukaiseen järjestykseen. Tällöin on päätettävä, tutkitaanko jakaumaa hyvinvoinnin (welfare) vai eriarvoisuuden (inequality) kannalta. Tässä tutkimuksessa ollaan kiinnostuneita eriarvoisuuden mittaamisesta, mutta koska sen voidaan ajatella olevan hyvinvoinnin osatekijä, joudutaan näitä käsittelemään yhdessä jakaumatarkastelujen teoriaa tarkasteltaessa.

2.1 Hyvinvoinnin mittaamisesta

Jakauman yhteiskunnallinen paremmuusvertailu perustuu jakauman tuottamaan hyvinvointiin (Dalton 1920). Atkinson (1970) esitti teoreettiset perusteet hyvinvointiin perustuvalle jakaumien eriarvoisuuden vertailulle.[3] Shorrocks (1983) laajensi käsittelyn koko hyvinvoinnin vertaamiseen, ja näihin esityksiin perustuu nykyinen tulojakauman hyvinvointi- ja eriarvoisuustutkimus.

Jakaumaan liittyvää hyvinvointia mitataan yhteiskunnan hyvinvointifunktiolla (social welfare function) W, joka tässä tutkimuksessa on määritelty jakauman tuottamana keskimääräisenä hyötynä

$$W = \int_{y_{min}}^{y_{max}} U(y)\, dF(y), \tag{1}$$

jossa U on hyötyfunktio ja F on tulojen y kertymäfunktio. Havaintoaineistoon perustuva tulojakauma ei ole jatkuva, ja siten hyödyn odotusarvo on luonnollista esittää muodossa

$$W(\mathbf{y}) = \frac{1}{N} \sum_{j=1}^{N} U(y_j). \tag{2}$$

Jos yhtälö (1) mielletään Stieltjes-integraaliksi, ovat edellä esitetyt yhtälöt yhtäpitäviä. Muoto (2) on tulkinnaltaan ilmeinen ja käytännön laskujen kannalta hyvä muoto. Useat teoreettiset esitykset suosivat kuitenkin ensimmäistä muotoa lisäten vielä jakaumalle jatkuvuusoletuksen. Tämä on perusteltua, kun mittausvirheille oletetaan jatkuva jakauma. Luvun 3 teoreettisessa osiossa käytetään molempia lähestymistapoja tarpeen mukaan.

Keskimääräisen hyödyn käyttöä voidaan helpoiten perustella ulkopuolisen tarkkailijan avulla. Ensimmäisen tulkinnan mukaan hyvinvoinnin määrittää eettinen tarkkailija, joka seuraa sivusta jakaumia ja laskee niille keskimääräisen hyödyn käyttäen itse päättämäänsä hyötyfunktiota. Toinen tulkinta on ”Veil of ignorance” -ajatuskoe, jossa tarkkailija on myös jakaumien populaatioiden ulkopuolella, mutta tietää joutuvansa satunnaisen yksilön paikalle valitsemaansa jakaumaan.[4] Käyttäytymisoletuksena on,

[3] Atkinsonia ennen ongelmaa tarkasteli Kolm, mutta hänen esityksensä on jäänyt suurelta osin vaille huomiota (Lambert 2001, 44).

[4] Henkilön sijainti jakaumassa ei tosin ole täysin sattumanvaraista, jos henkilö tuntee omat ominaisuutensa ja ahkeruutensa.

että tällöin hän laskee odotetun hyödyn jakaumille ja valitsee sen, jossa se saa suurimman arvonsa. Hyötyfunktio on tässä tapauksessa ymmärrettävä von Neumann – Morgenstern -hyötyfunktioksi.[5] (Lambert 2001, jakso 4.1.)

Keskimääräinen hyöty voidaan perustella myös populaatioon kuuluvien yksilöiden kokemista hyödyistä. Tällöin jakaumaan liittyvää hyvinvointia kuvaa yhteiskunnan hyvinvointifunktionaali W_U, joka määrää funktion W^* yhtälöllä

$$W^*(\mathbf{y}) = W_U\left(\mathbf{U}(\mathbf{y})\right). \tag{3}$$

Kun funktiolle W^* asetetaan joukko ehtoja (symmetrisyys, yksilöiden itsekeskeisyys ja W_U:n additiivinen separoituvuus), päädytään kardinaaliseen hyötyteoriaan, klassiseen utilitarismiin mutta myös haluttuun hyvinvointifunktion muotoon eli yhtälöön (2) (Lambert 2001, jakso 4.2). Koska kardinaalinen hyvinvointiteoria on ongelmallinen, lienee paras perustella keskimääräiseen hyötyyn perustuva hyvinvointivertailu ulkopuolisen tarkkailijan avulla.

Ei ole kuitenkaan ilmeistä, että keskimääräinen hyöty olisi hedelmällinen lähestymistapa eriarvoisuustarkasteluihin. Ovathan kiinnostuksen kohteena juuri yksilöiden hyvinvointitasojen erot eikä niiden summa tai keskiarvo. Kun yksilöiden hyötyfunktiot ovat identtiset ja konkaavit, utilitaristisen hyvinvointifunktion maksimoiva tulonjako on annetulla tulojen keskiarvolla tulojen tasajako. Siten utilitaristisen hyvinvointifunktion käyttöä voidaan tässä yhteydessä perustella, joskaan esimerkiksi Sen (1973, 15–18) ei pidä utilitarismia riittävänä lähtökohtana eriarvoisuuden tarkastelemiseen.

Kuten edellä mainittiin, hyötyfunktio U oletetaan tyypillisesti aidosti kasvavaksi ja konkaaviksi ($U' > 0$, $U'' < 0$). Aidosti kasvavan funktion oletus perustuu ajatukseen, että tulosta saatu hyöty ei saturoidu, vaan jokainen tienattu lisäeuro lisää yksilön hyötyä. Konkaavisuus puolestaan perustuu käsitykseen, jonka mukaan matalatuloinen hyötyy lisätulosta korkeatuloista enemmän. Tämä voidaan käsittää tarkkailijan haluksi välttää eriarvoisuutta ja ”Veil of ignorance” -ajattelussa riskin kaihtamiseksi. Nämä oletukset tehdään hyötyfunktiolle tässäkin tutkimuksessa.

Lienee syytä muistuttaa, että hyvinvoinnilla tarkoitetaan tässä tutkimuksessa yksittäisen tulomuuttujan jakauman sisältämän yhteiskunnallisen hyödyn tarkastelua, kun yhteiskunnallinen hyöty on määritelty yhtälöllä (2). Todellinen yhteiskunnallinen hyvinvointi koostuu tulojen lisäksi monista muista taloudellisista ja ei-taloudellisista tekijöistä.

2.2 Suhteellinen ja absoluuttinen eriarvoisuus

Kun jakauman hyvinvoinnin kriteerinä on yhtälö (2), jossa $U(y)' > 0$ ja $U(y)'' < 0$, on ilmeistä, että kasvava tulotaso y nostaa hyvinvointia. Mutta koska suurituloiset hyötyvät lisätuloista matalatuloisia vähemmän, on yhtä ilmeistä, että tulon tasainen jakautuminen lisää myös hyvinvointia. Lähtökohtana oleva hyötyfunktio maksimoituu

[5] Koska odotetun hyödyn maksimointiin perustuvaan käyttäytymismallintamiseen liittyy ongelmia (esim. Kahneman ja Tversky 1979), voisi kuvitella, että odotettuun hyötyyn perustuvaan luokitteluun voi myös liittyä ongelmia.

kaikkien yksilöiden saadessa saman tulon $y_j = \overline{y}$, kun rajoitteena on tulojen keskiarvon vakioisuus. Tulojen poikkeaminen tästä keskitulosta aiheuttaa hyötyfunktion arvojen hajontaa, jota kutsutaan johdannossa esitetyllä tavalla (taloudelliseksi) eriarvoisuudeksi.

Tuloeroja ja siten eriarvoisuutta voidaan tarkastella kahdella eri tavalla: Voidaan tutkia tulojen suhteellista hajontaa, jossa mielenkiinnon kohteena on suure

$$\frac{y_i - y_j}{\overline{y}}, \tag{4}$$

jossa $\overline{y}$ on tulojen keskiarvo. Tällöin tulojen kertominen positiivisella vakiolla ei muuta tutkittua suuretta, mutta positiivisen vakion lisääminen laskee sitä. Toisaalta voidaan tutkia suuretta

$$y_i - y_j, \tag{5}$$

jolloin kertominen positiivisella vakiolla kasvattaa suuretta, mutta sen lisääminen ei muuta suuretta. Ensimmäinen suure vastaa suhteellisen ja jälkimmäinen absoluuttisen eriarvoisuuden mielenkiinnon kohdetta. Tulopoliittisten neuvottelujen prosenttilinja vastaa siten suhteellisen eriarvoisuuden muuttumattomuutta, pennilinja absoluuttisen.

Usein, ja niin myös tässä tutkimuksessa, ollaan kiinnostuneita suhteellisesta eriarvoisuudesta. Näin eriarvoisuus saadaan riippumattomaksi hyvinvoinnin toisesta tekijästä keskiarvosta. Ei ole kuitenkaan täysin varmaa, kuvaako se tuloeriarvoisuuden kaikkia puolia. Voidaan esimerkiksi ajatella, että yhteiskunnassa vallitsee kulutusnormi, jonka alapuolelle jääminen merkitsee absoluuttista köyhyyttä. Tällöin on mahdollista, että suhteellisten tuloerojen pysyessä vakiona, voi alle kulutusnormin jäävien kotitalouksien osuus kasvaa. Lisäksi absoluuttinen etäisyys tähän normiin kasvaa myös. Tämä kasvu merkitsee eriarvoisuuden kasvua, vaikka emme sitä suhteellisen eriarvoisuuden mitoilla kykenekään havaitsemaan. Tällainen problematiikka liittynee kuitenkin enemmän köyhyystutkimukseen. Edellä ja jatkossa puhuttaessa eriarvoisuudesta tarkoitetaan suhteellista eriarvoisuutta, jollei toisin mainita.

Lopuksi on syytä huomauttaa, että eriarvoisuus ja tulojen keskiarvo voivat olla kytkeytyneitä toisiinsa. Voidaan nimittäin perustellusti väittää, että tuloerojen tasaantuminen vähentää esimerkiksi työnteon ja koulutuksen kannustinvaikutuksia. Tällöin vähenevä eriarvoisuus voi vähentää myös tulojen keskiarvoa, koska kotitalouksilla ei ole kannusteita ahkeroida ja yrittää.

2.3 Lorenz-käyriin perustuva vertailu

Atkinson (1970) osoitti, että jos jakaumien keskiarvot ovat samat, voidaan jakaumien hyvinvointivertailu suorittaa Lorenz-käyriin perustuen. Edellisen jakson periaatteiden mukaan kaikki konkaavit hyötyfunktiot luokittelevat jakauman F G:tä paremmaksi, jos jakauman F Lorenz-käyrä on jossain G:n käyrän yläpuolella, muttei missään alapuolella (F dominoi G:tä). Dominanssi-tulosta on myöhemmin yleistetty. Tässä jaksossa on lyhyt esittely aiheeseen, mutta aihetta on käsitelty laajasti muualla (esim. Aura 1996).

2.3.1 Lorenz-dominanssi

Lorenz-käyrä on yleisesti käytetty tulojakauman tasaisuutta kuvaava käyrä. Se kuvaa tulonsaajaosuuden p_i saaman osuuden Φ_i kokonaistuloista, kun tulonsaajat on järjestetty tulojen mukaan pienimmästä suurimpaan. Siten Lorenz-käyrä $L(p)$ voidaan yleisesti määritellä tulojen y kertymäfunktion $F(y)$ käänteisfunktion

$$F^{-1}(p) = \inf_y \{y : F(y) \geq p\} \tag{6}$$

avulla muodossa

$$L(p; y) = \mu^{-1} \int_0^p F^{-1}(x)dx, \quad 0 \leq p \leq 1, \tag{7}$$

jossa μ on tulon keskiarvo (Gastwirth 1971). Näin siksi, että yhtälö (7) kuvaa tulo-osuuden kertymää ja kertymäfunktion käänteisfunktio huolehtii siitä, että tulo-osuus muodostuu oikein suhteessa tulonsaajien kertymään.

Edellä oleva muotoilu pätee myös diskreetille jakaumalle, mutta käytännössä operationaalisessa muodossaan Lorenz-käyrä on helpoin esittää kahden yhtälön avulla, eli yksi yhtälö kummallekin kertymälle. Tällöin p_i voidaan esittää, kun erisuuruisia tuloja on n kappaletta, muodossa

$$p_i = \frac{1}{N} \sum_{j=1}^{n} f_j \mathrm{T}(y_j \leq y_i) = \frac{1}{N} \sum_{j=1}^{i} f_j, \tag{8}$$

jossa f_j on tulon j frekvenssi, $N = \sum_k f_k$ ja T on totuusfunktio, joka saa arvon 1, kun argumentti on tosi ja muulloin arvon 0. Yhtälön ensimmäinen yhtäsuuruus on määritelmä ja toinen yhtäsuuruus pätee, kun havainnot on järjestetty pienimmästä suurimpaan ja indeksoitu sopivasti uudelleen. Vastaavalla tavalla havainnot uudelleen järjestämällä saadaan tulo-osuuden kertymälle muoto

$$\Phi_i = \Phi(p_i) = \frac{1}{N\overline{y}} \sum_{j=1}^{n} f_j y_j \mathrm{T}(y_j \leq y_i) = \sum_{j=1}^{i} w_j \frac{y_j}{\overline{y}}, \tag{9}$$

jossa w_j on tulon y_j suhteellinen frekvenssi f_j/N ja $\overline{y} = \sum_k w_k y_k$.[6] Lorenz-käyrä voidaan siten esittää muodossa

$$L(p_i; \mathbf{y}) = \sum_{j=1}^{i} w_j \, r_j, \tag{10}$$

jossa $r_j = y_j/\overline{y}$ ja $L(0; \mathbf{y}) = 0$. Muuttujaa r_j kutsutaan jatkossa skaalatuksi tuloksi. Vektori $\mathbf{y}$ ilmaisee, minkä muuttujan jakaumaa tarkastellaan, mutta on myös oikeastaan argumentti, sillä jos jokin vektorin komponentti muuttuu, muuttuvat myös jotkin Lorenz-ordinaatan arvot. Ääripäissään Lorenz-käyrä saa arvot $L(0; \mathbf{y}) = 0$ ja $L(1; \mathbf{y}) = 1$. Jatkuva muoto $L(p; \mathbf{y})$ määritellään siten, että diskreettien argumentin arvojen p_i ja p_{i+1} välillä Lorenz-käyrä saa arvot

$$L(p_i + \tau \Delta p_i; \mathbf{y}) = (1 - \tau) L(p_i; \mathbf{y}) + \tau L(p_{i+1}; \mathbf{y}), \quad \forall \, \tau \in [0, 1], \tag{11}$$

[6] Lienee todennäköistä, että populaatiossa ei ole monta täsmälleen samaa tuloa (nollatuloa lukuun ottamatta). Onkin luultavaa, että monessa tapauksessa $w_j = 1/N$ ja tällöin $n \approx N$. Yllä olevat muodot pätevät luonnollisesti myös tällöin.

jossa $\Delta p_i = p_{i+1} - p_i$, eli Lorenz-käyrän diskreetit pisteet yhdistetään suorilla. Otosaineistoa käsiteltäessä tällöin implisiittisesti oletetaan, että kahden havainnon välillä ei ole tuloeroja, vaan kaikki välissä olevat saavat samaa tuloa kuin havainnoista suurempi. Tätä ongelmaa käsitellään myöhemmin.

Edellä kuvattu määritelmä sallii periaatteessa negatiiviset tulot, kunhan tulojen keskiarvo on positiivinen. Teoreettiset tarkastelut koskevat kuitenkin tyypillisesti tilannetta, jossa tulot ovat ei-negatiivisia tai jopa positiivisia. Tällöin Lorenz-käyrä saa vain ei-negatiivisia arvoja. Huomattavaa on se, että Lorenz-käyrä on skaalainvariantti, eli tulovektori voidaan kertoa positiivisella vakiolla, ilman että Lorenz-käyrä muuttuisi. Tämän ominaisuuden vuoksi käyrä soveltuu suhteellisen eriarvoisuuden tutkimiseen.

Kuten edellä mainittiin, Atkinson (1970) osoitti, että eriarvoisuutta välttävä tarkkailija preferoi sitä jakaumaa, joka Lorenz-dominoi muita vertailtavia jakaumia, kun jakaumien keskiarvot ovat samat. Tällöin Lorenz-dominoivaan jakaumaan liittyy korkein hyvinvointifunktion arvo. Tämä johtuu siitä, että keskiarvojen ollessa samat, on dominoivalla jakaumalla vähiten suhteellista eriarvoisuutta. Jos keskiarvot eivät ole samat, Lorenz-käyrä ei sisällä kaikkea hyvinvointia määrittävää tietoa. Voidaan kuitenkin osoittaa, että Lorenz-dominoivassa jakaumassa on yhä pienin suhteellinen eriarvoisuus, vaikka hyvinvointeja ei voidakaan Lorenz-käyrin järjestää.

Dasgupta et al. (1973) osoittivat, että Lorenz-dominanssi vastaa korkeampaa hyvinvointia yleisemmälläkin hyvinvointifunktionaalilla kuin mitä edellisessä jaksossa vaadittiin. Voidaan nimittäin osoittaa, että merkitsemällä

$$W = W_U\left(\mathbf{U}(\mathbf{y})\right) \equiv W^*(\mathbf{y}), \tag{12}$$

funktio W^* järjestää jakaumat samoin kuin Lorenz-käyrä, jos W^* on S-konkaavi. Funktio f on S- eli Schur-konkaavi, jos pätee $f(\mathbf{Bx}) \geq f(\mathbf{x})$, kun $\mathbf{B}$ on bistokastinen matriisi. Bistokastisen matriisin alkiot ovat ei-negatiivisia ja kaikki rivit ja sarakkeet summautuvat ykköseen, $b_{ij} \in [0, 1]$. Bistokastisen matriisin kuvaama tulonsiirto on keskiarvon säilyttävä ja tulojakaumaa ei-leventävä.[7] S-konkaavi hyvinvointifunktio kuvaa siis eriarvoisuutta kaihtavaa yhteiskuntaa. Lorenz-dominanssiin liittyvä hyvinvointivertailu voidaan siis perustella yhteiskunnan sisäisin tarkasteluin tarvitsematta oletusta hyötyjen additiivisesta separoituvuudesta.

2.3.2 Yleistetty Lorenz-dominanssi

Shorrocks (1983) yleisti Atkinsonin tuloksen koskemaan jakaumia, joilla on eri keskiarvot. Tähän tarvittiin Lorenz-käyrän yleistys, ns. yleistetty Lorenz-käyrä (GL), joka on määritelty yhtälöllä

$$GL(p_i;\mathbf{y}) = \overline{y}\, L(p_i;\mathbf{y}) \tag{13}$$

eli

$$GL(p_i;\mathbf{y}) = \sum_{j=1}^{i} w_j\, y_j. \tag{14}$$

[7] Rivisumma pitää huolen, että kombinaatio on konveksi. Sarakesumma puolestaan varmistaa, että kukin tulo jaetaan kokonaan eli tulosumma säilyy.

Lisäksi jatkuvassa esityksessä $GL(p;\mathbf{y})$ diskreetit pisteet yhdistetään suoralla, kuten tavallisen Lorenz-käyrän tapauksessakin tehtiin. Shorrocks osoitti, että on voimassa lause

$$W^*(\mathbf{y}) \geq W^*(\mathbf{y}') \ \ \forall\, W^* \in \mathcal{W} \Leftrightarrow GL(p;\mathbf{y}) \geq GL(p;\mathbf{y}') \ \ \forall\, p, \tag{15}$$

eli jos yleistetty Lorenz-käyrä dominoi toista, on siihen liittyvä hyvinvointi dominoitua jakaumaa suurempi, kun hyvinvointifunktio kuuluu ei-vähenevien S-konkaavien funktioiden ($\mathcal{W}$) joukkoon.

Yleistetyn Lorenz-käyrän yksikkö on mitatun tulon yksikkö, joten eri maiden sekä eri ajan hetkinä mitattujen jakaumien vertailu on tavallista Lorenz-käyrää haastavampaa, sillä tuloyksiköt täytyy normittaa jollain sopivalla tavalla.

2.3.3 Absoluuttinen Lorenz-dominanssi

Siinä missä tavallinen Lorenz-käyrä soveltuu suhteellisen eriarvoisuuden tutkimiseen, soveltuu Moyesin (1987) esittämä absoluuttinen Lorenz-käyrä (AL) absoluuttisen eriarvoisuuden mittaamiseen. Absoluuttinen Lorenz-käyrä on määritelty yhtälöllä

$$AL(p_i;\mathbf{y}) = GL(p_i;\mathbf{y}^*), \tag{16}$$

jossa $\mathbf{y}^* = \mathbf{y} - \mathbf{1}\overline{y}$ eli voidaan kirjoittaa

$$AL(p_i;\mathbf{y}) = \sum_{j=1}^{i} w_j\,(y_j - \overline{y}). \tag{17}$$

Kun tavallinen Lorenz-käyrä on invariantti tulojen suhteellisen muutoksen suhteen, on absoluuttinen Lorenz-käyrä invariantti additiivisen muutoksen suhteen, mikä käy suoraan ilmi yhtälöstä (17). Absoluuttisen Lorenz-käyrän yksikkö on mitatun tulon yksikkö, joten sen käyttöön pätee sama huomautus kuin yleistettyyn Lorenz-käyrään edellä.

2.4 Leikkaavat Lorenz-käyrät

Lorenz-käyriin perustuvat menetelmät mahdollistavat vain jakaumien osittaisen järjestämisen eli kaikkia jakaumia ei saada paremmuusjärjestykseen. Dominanssia ei nimittäin esiinny, jos Lorenz-käyrät leikkaavat toisensa. Jotta jakaumavertailu voitaisiin silloinkin tehdä yhtälön (2) mielessä, täytyy tutkijan tehdä lisäoletuksia hyötyfunktion U muodosta. Näin menetetään edellä kuvattujen menetelmien yleisyys. Hyvinvointivertailuja voidaan kuitenkin joissain tapauksissa vielä tehdä pienin lisäoletuksin ja lisälaskelmin yleistettyjen Lorenz-käyrien avulla (Lambert 2001, jakso 3.4). Jotta jakaumat saataisiin täydelliseen järjestykseen, tarvitaan kuitenkin voimakkaita lisäoletuksia. Tutkijan on nimittäin otettava kantaa siihen, kuinka tulojen suuruutta ja eriarvoisuutta painotetaan hyvinvointitarkasteluissa.

Jakauman sisältämä hyvinvointi voidaan yrittää tiivistää kahteen tunnuslukuun: keskiarvoon ja suhteellisen eriarvoisuuden indeksiin eli eriarvoisuusmittaan. Voidaan siis muodollisesti määritellä hyvinvointifunktio

$$W = W^*(\mathbf{y}) = V\left(\overline{y}, I(\mathbf{y})\right), \tag{18}$$

jossa $\overline{y}$ on keskiarvo ja $I(\mathbf{y})$ eriarvoisuusmitta (esim. Lambert 2001, luku 5). Yhtälössä esiintyvä eriarvoisuusmitta $I(\mathbf{y})$ on suure, jonka avulla pyritään tiivistämään jakauman sisältämä eriarvoisuus yhdeksi luvuksi. Mitta on määritelty siten, että vähenevä eriarvoisuus pienentää mitan arvoa. Siten funktiolle V päteekin tyypillisesti

$$\frac{\partial V}{\partial \overline{y}} > 0 \quad \text{ja} \quad \frac{\partial V}{\partial I} < 0. \tag{19}$$

Jaksossa 2.2 todettiin, että tulojen keskiarvo ja eriarvoisuus voivat olla kytköksissä siten, että aleneva eriarvoisuus supistaa keskimääräistä tuloa, koska kannustimet ponnistella vähenevät. Voidaan siis olettaa, että ainakin pienillä eriarvoisuusmitan arvoilla pätee

$$\frac{\partial \overline{y}}{\partial I} > 0. \tag{20}$$

Osittaisderivaatta voi olla myös negatiivinen, jos eriarvoisuus on jo niin suurta, että sen lisääminen haittaa yhteiskuntarauhaa ja aiheuttaa siten tuotannontason laskua.

Edellisen lisäksi on syytä huomioida se, että tuloerot voivat olla liian pienet jo oikeudenmukaisuudenkin perusteella. Tällöin erilaisia kohdellaan sosiaaliseen normiin nähden liian samanlaisesti, mikä johtaa siihen, että eriarvoisuuden lisääminen lisää hyvinvointia, vaikka keskimääräiset tulot pysyisivätkin vakioina. Riittävän pienillä eriarvoisuusmitan I arvoilla pätee siis

$$\frac{\partial V}{\partial I} > 0. \tag{21}$$

Siten optimaalinen hyvinvointi saavutetaan nollasta poikkeavan eriarvoisuuden vallitessa. Muodollisesti optimaalisen eriarvoisuusindeksin arvo saadaan ehdosta

$$\frac{dV}{dI} = \frac{\partial V}{\partial I} + \frac{\partial V}{\partial \overline{y}} \frac{\partial \overline{y}}{\partial I} = 0, \tag{22}$$

jossa oletettavasti $V_{\overline{y}} > 0$ kaikilla y:n arvoilla.

Tarkastellaan optimiehdon derivaatan arvoa kolmessa eri yhteiskunnallisessa tilassa: Tilassa 1 tuloerot ovat erittäin pienet ja alle sosiaalisen normin. Tällöin $V_I > 0$ ja $\overline{y}_I > 0$ eli eriarvoisuuden kasvu lisää hyvinvointia suoraan sekä lisääntyvän tuotannon (tulon) kautta. Tällöin derivaatta on positiivinen. Tilassa 2 tuloerot ovat maltilliset, mutta sosiaalista normia suuremmat. Tällöin $V_I < 0$ ja $\overline{y}_I > 0$ eli eriarvoisuuden kasvu vähentää suoraan hyvinvointia, mutta lisää sitä yhä tuotannon kautta ja siten derivaatan merkki on tuntematon. Tilassa 3 tuloerot ovat valtaisat ja aiheuttavat jo levottomuutta. Tällöin $V_I < 0$ ja $\overline{y}_I < 0$ eli eriarvoisuuden kasvu vähentää hyvinvointia sekä suoraan, että vähentyvän tuotannon kautta. Tällöin derivaatta on negatiivinen. Tästä seuraa, että yhteiskunnallisesti optimaalinen eriarvoisuusmitan arvo löytyy tilasta 2, jolloin eriarvoisuus on yli sosiaalisen normin, mutta sillä on negatiivisia kannustinvaikutuksia. Niin eriarvoisuuden sosiaaliseen normiin kuin optimaalisen suuruuteenkaan ei tässä tutkimuksessa oteta tämän tarkemmin kantaa, vaan oletuksena on, että Suomi on luultavasti tilassa 2. Se, ovatko tuloerot optimaalisen tason ylä- vai alapuolella, vaatii sosiaalisten normien lähempää tarkastelua.

Toisin kuin Lorenz-käyrillä, funktion V avulla voidaan jakaumat asettaa täydellisesti järjestykseen, mutta kuten edellä todettiin, joudumme tällöin ottamaan kantaa siihen, miten tehokkuutta (tulotasoa) ja eriarvoisuutta funktiossa V painotetaan. Tässä työssä ei kuitenkaan tarkastella jakaumien hyvinvointieroja vaan tuloeroja sekä suhteellista eriarvoisuutta. Nämä tarkastelut tehdään edellä kuvatun eriarvoisuusmitan käsitteen avulla ja havaittu eriarvoisuus kytkeytyy hyvinvointiin, joka on jakaumien hyvyyden vertailukriteeri, edellä kuvatulla tavalla.

Siirryttäessä tarkastelemaan tuloeroja ja eriarvoisuutta ei tutkijan tarvitse määritellä hyvinvointifunktiota V, mutta hänen täytyy suorittaa eriarvoisuusmitan I valinta. Mitan valinta on oma ongelmansa, sillä eriarvoisuusmitta voidaan määritellä monella tavalla riippuen siitä, kuinka tulojakauman eri osia halutaan painottaa. Eriarvoisuusmittoja käsitellään tarkemmin seuraavassa luvussa.

3 Eriarvoisuusmitat

Edellisessä luvussa kuvattu Lorenz-dominansseihin perustuva jakaumavertailu on yleisyytensä vuoksi teoreettisesti miellyttävä. Menetelmät ovat riippumattomia tutkijan normatiivisista valinnoista ja lisäksi menetelminä täysin parametrittomia. Toisinaan Lorenz-dominanssi ei kuitenkaan tutkijalle riitä. Tyypillinen ongelma on se, että Lorenz-käyrät leikkaavat eikä Lorenz-dominanssia esiinny. Toisinaan taas halutaan saada numeerista tietoa eriarvoisuuden muutoksista. Näissä tapauksissa tutkijan on turvauduttava suhteellisen eriarvoisuuden indekseihin (jatkossa eriarvoisuusmittoihin).[8]

3.1 Määrittelevät ominaisuudet

Jotta jakaumat järjestettäisiin eriarvoisuuden mielessä oikein, on määritettävä kriteerit, milloin kahta jakaumaa verrattaessa jakaumassa on toista enemmän (tai yhtä paljon) eriarvoisuutta. Jakauman tunnusluku on kelvollinen kuvaamaan jakaumassa esiintyviä tuloeroja ja eriarvoisuutta, jos se noudattaa näitä kriteereitä. Koska tutkimme eriarvoisuusmittoja, on mitan arvon oltava sitä suurempi, mitä enemmän eriarvoisuutta jakaumassa ilmenee.

Nämä eriarvoisuuskriteerit ovat siis mitalta vaadittavia ominaisuuksia, jotka määrittävät sen antamien arvojen suhteet eri jakaumien välillä. Tällaisia ominaisuuksia voidaan luonnollisesti määrittää lukuisia, sillä nämä ominaisuudet määrittävät sen, mitä tässä yhteydessä käsitämme eriarvoisuudella. Yleisesti vaaditut kolme ominaisuutta ovat:

1. Symmetria: Havaintoyksikköjärjestyksellä ei ole merkitystä eriarvoisuuden määrään. Mitan laskennassa mahdollisesti tarvittavat havaintojen uudelleen järjestelyt on perustettava tutkittujen tulojen arvoihin.

2. Populaatioperiaate (Dalton 1920): Populaation koko ei saa vaikuttaa eriarvoisuusindeksin arvoon. Populaatioperiaatteen täyttävän mitan arvo ei siis muutu, kun n samanlaista populaatiota yhdistetään.

3. Siirtoperiaate (Dalton 1920): Jos muiden havaintoyksiköiden tulot pysyvät ennallaan ja tuloparia $(y, y + \Delta)$, $\Delta > 0$ muutetaan tulosiirrolla suurituloiselta pienituloiselle siten, että uusi tulopari on $(y + c\Delta, y + (1 - c)\Delta)$ ja $c \in (0, 1/2)$ eli tuloparien keskinäinen järjestys ei muutu. Tällöin eriarvoisuus vähenee ja siten periaatteen täyttävän eriarvoisuusmitan arvo pienenee. Tärkeää siirrossa on siis se, että tulojen summa säilyy, siirto on suurituloiselta pienituloiselle ja että kotitalouksien tulojen suuruusjärjestys ei muutu.[9] Jos jakauma F voidaan muodostaa jakaumasta G edellä kuvatuin tulonsiirroin, jakauma F Lorenz-dominoi

[8] Eriarvoisuusmittojen käytön ideaa sekä erilaisia mittoja esitelevät esimerkiksi Sen (1973, luku 2) ja Cowell (1995, luku 3), joskin laskukaavat on esitetty tapaukselle, jossa kaikilla havainnoilla on sama paino ($1/N$), eivätkä siten sovellu suoraan otosaineistojen tutkimiseen. Cowellin (2000) artikkeli on hyvä esimerkki teknisemmästä esityksestä.

[9] Tämä esitetään usein toisinpäin, eli siirto tapahtuu pienituloiselta suurituloisemmalle, jolloin eriarvoisuus luonnollisesti kasvaa (esim. Cowell 2000). Tässä käytetty muotoilu on peräisin Daltonin (1920) artikkelista.

jakaumaa G. Lisäksi jos eriarvoisuusmitta toteuttaa siirtoperiaatteen, antaa se aina pienemmän arvon jakaumalle F kuin jakaumalle G (Cowell 1995, 55).

Näistä ominaisuuksista kaksi ensimmäistä määrittelee, millä ehdoin jakaumissa on sama määrä eriarvoisuutta ja kolmas sen, milloin eriarvoisuus on jakaumassa toista vähäisempää. Siirtoperiaate on suora seuraus ajatuksesta, jonka mukaan jakaumaan liittyvä hyvinvointi voidaan yhtälön (18) mukaisesti tiivistää tunnuslukuihin keskiarvo ja eriarvoisuusindeksi. Tällöin keskiarvon säilyttävä siirto muuttaa vain eriarvoisuusindeksiä ja koska laskeva indeksi nostaa hyötyä, on indeksin laskettava siirron seurauksena. Lisäksi koska W^* on argumenttiensa suhteen symmetrinen, kuten keskiarvokin, on sitä oltava myös eriarvoisuusindeksin (Lambert 2001).

Näiden lisäksi tarvitaan vielä neljäs ominaisuus, joka määrittelee eriarvoisuuskäsitteeksi suhteellisen eriarvoisuuden.

4 Skaalainvarianssi: Tunnusluku ei riipu siitä, missä yksiköissä tulo on mitattu. Eli kaikkien tulojen kertominen positiivisella vakiolla a ei muuta ehdon toteuttavan mitan arvoa. Kyseessä on siis invarinssi suhteellisen lisäyksen suhteen.

Ominaisuus määrittää mitan suhteellisen eriarvoisuuden mitaksi. Absoluuttisen eriarvoisuuden tapauksessa invarianssi olisi absoluuttisen lisäyksen suhteen. Jos eriarvoisuusmitalta vaaditaan edellä esitellyt neljä ominaisuutta, voidaan yleisesti käytetyistä mitoista rajoittua kolmeen tapaukseen: (laajennettuun) Gini-kertoimeen, Atkinsonin indeksiin sekä yleistettyyn entropiamittaan. Näiden ominaisuuksia tarkastellaan lähemmin seuraavassa jaksossa.

3.2 Eräitä eriarvoisuusmittoja

Eriarvoisuusmittoja on lukuisia ja ne täyttävät edellä kuvatut kriteerit vaihtelevalla menestyksellä. Tässä jaksossa käsitellään kolmea sellaista mittaperhettä, jotka toteuttavat edellä esitetyt neljä ominaisuutta. Nämä kolme mittaperhettä ovat myös eriarvoisuusmitoista yleisimmin käytetyt.

3.2.1 Gini-kerroin

Gini-kerroin on kenties yleisimmin käytetty eriarvoisuusmitta (Gini 1912, 1914 ja 1921). Sen graafinen tulkinta on Lorenz-käyrän ja tasajakosuoran välinen ala (ns. Lorenz-ala) kerrottuna kahdella eli yhtälönä

$$G = 2\int_0^1 [p - L(p)]\,\mathrm{d}p, \tag{23}$$

joka voidaan yhtäpitävästi kirjoittaa muodossa

$$G = 1 - 2\int_0^1 L(p)\,\mathrm{d}p, \tag{24}$$

jossa $L(p)$ on jakauman Lorenz-käyrä (ks. yhtälö (10)). Geometrisen tulkinnan perusteella on ilmeistä, että Gini-kerroin voi, ei-negatiivisten tulojen tapauksessa, saada arvoja vain nollan ja yhden väliltä. Gini-kertoimen arvo nolla tarkoittaa tulojen tasajakoa; arvoa yksi vastaa tilanne, jossa infinitesimaalinen osa populaatiosta saa kaiken jaettavissa olevan tulon. Edellisten lisäksi osittaisintegroimalla yhtälöstä (23) Gini-kertoimelle saadaan esitys

$$G = 2\int_0^1 p\,[r(p) - 1]\,\mathrm{d}p, \tag{25}$$

josta lyhyellä manipulaatiolla päädytään muotoon

$$G = \int_0^1 [2p - 1]\,r(p)\mathrm{d}p, \tag{26}$$

joissa $r(p) = y(p)/\mu$ eli skaalattu tulo.

Kuten yltä huomataan, voidaan Gini-kerroin laskea monella ekvivalentilla tavalla ja nämä eri esitysmuodot tuovat esiin Gini-kertoimen eri tulkinnat. Sekä erilaisia esitysmuotoja että erilaisia tulkintoja ovat esittäneet erittäin kattavasti Nygård ja Sandström (1981, 240–244 ja taulu 7.2). Edellisten lisäksi Gini-kerroin voidaan määritellä tulon y ja tulon kertymäfunktion $F(y)$ avulla. Näihin muotoihin ei tässä työssä perehdytä, vaan siirrytään nyt tarkastelemaan Gini-kertoimen operationaalisia esityksiä. Ei nimittäin ole täysin suoraviivaista soveltaa integraalimuotoja havaintoaineistoon.

Pinta-ala-määritelmästä lähtevä ratkaisu pohjautuu Lorenz-käyrän integrointiin. Diskreetissä tapauksessa integrointi johtaa yhtälöön (ks. Liite A)

$$G = 1 - \sum_{j=1}^{n} w_j(\Phi_{j-1} + \Phi_j), \tag{27}$$

jossa Φ on tulo-osuuden kertymä ($\Phi_0 = 0$ ja $\Phi_n = 1$) ja w_j on suhteellinen frekvenssi f_i/N, jossa $N = \sum_k f_k$ eli populaation jäsenten lukumäärä. Käyttäen yhtälöön (27) tulo-osuuden kertymän Φ määritelmää (9) päästään lyhyellä laskulla (ks. Liite A) muotoon

$$G = 2\left(\sum_{j=1}^{n} p_j w_j r_j - \frac{1}{2}\sum_{j=1}^{n} w_j^2 r_j\right) - 1, \tag{28}$$

jossa $r_j = y_j/\overline{y}$ ja p_j on tulonsaajien kertymä $\sum_{k=1}^{j} w_k$, eli tuloa y_j matalatuloisempien kotitalouksien osuus aineistosta. Saatu yhtälö vastaa edellä esitettyä yhtälöä (25), kun huomataan, että $2\int_0^1 p\,\mathrm{d}p = 1$.

Toinen näkökulma määritellä Gini-kerroin on lähteä liikkeelle tuloparien keskimääräisistä absoluuttisista eroista. Tutkittaessa tuloparia (x, y) Gini-kertoimen määritelmä on

$$G = \frac{E|x - y|}{2\mu}, \tag{29}$$

jossa E on odotusarvo-operaattori. Gini on siis myös puolet kahden tulonsaajan välisestä keskimääräisestä absoluuttisesta tuloerosta jaettuna keskimääräisellä tulolla. Käytännön laskuissa tämä johtaa muotoon

$$G = \frac{1}{2\overline{y}}\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j |y_i - y_j|, \tag{30}$$

jossa oleva kaksoissumma on varsin hidas tapa laskea mitan arvo. Tästä vaihtoehtoisesta määritelmästä voidaan kuitenkin helposti johtaa lisää esitysmuotoja Gini-kertoimelle (esim. Dorfman 1979).

Nygård ja Sandström (1981) sekä Lerman ja Yitzhaki (1984) esittivät Gini-kertoimelle muodon

$$G = 2\,\mathrm{cov}(r, p), \tag{31}$$

eli Gini-kerroin on skaalatun tulon $r = y/\overline{y}$ ja kotitalouksien kertymän $p = F(y)$ kovarianssi (ks. yhtälöt (25) ja (26)). Yhtälön avulla Gini-kerroin voidaan laskea helposti kaikilla ohjelmistoilla, joissa on vektorin järjestämis- ja kovarianssiohjelmat.

Edellä mainittiin, että Gini-kerroin saa ei-negatiivisten tulojen tapauksessa arvoja nollan ja yhden väliltä. Edellä esitetyt esitysmuodot ovat kuitenkin voimassa myös, jos joukossa on negatiivisia tuloja, kunhan tulojen keskiarvo on positiivinen. Tällöin Lorenz-käyrä ei ole kaikkialla kasvava, mutta se on edelleen konveksi. Jos negatiivisia tuloja on paljon, voi Gini-kerroin saada ykköstä suurempia arvoja. Tällöin kertoimen arvon tulkinta vaikeutuu. Tyypillisesti vuosituloaineistoissa on myös negatiivisia tulohavaintoja, mutta niiden määrä on vähäinen ja siten Gini-kerroin saa arvoja normaalilta vaihteluväliltään.

Gini-kertoimen hyvinvointiteoreettinen kompastuskivi on se, että yhtälön (3) mukaisia derivoituvia funktioita W_U ja $\mathbf{U}$, jotka järjestäisivät jakaumat samalla tavalla kuin Gini-kerroin, ei ole olemassa[10] (Lambert 1985). Samoin jos hyvinvointifunktio on aidosti kvasi-konkaavi yksilöiden tulojen suhteen, ei Gini-kerroin pysty kuvaamaan tällaista järjestystä (Dasgupta et al. 1973). Siten yksilötason hyvinvointitarkastelut eivät ole suoraviivaisia Gini-kertoimen tapauksessa. Hyvinvointitarkastelut voidaan kuitenkin tehdä yhtälön (18) tapaan, jolloin yhteiskunnan hyvinvointifunktio on muotoa

$$V\,(\overline{y}, G)) = \overline{y}(1 - G), \tag{32}$$

jonka tulkintoja on koonnut Lambert (2001, 122–125).

3.2.2 Laajennettu Gini-kerroin

Edellä esitettyä Gini-kerrointa voidaan yleistää monin eri tavoin. Nygård ja Sandström (1981, 211) esittivät ns. lineaarisen eriarvoisuusmittaan (Mehran 1976)

$$I_{LIN} = \int_0^1 W(p)\,[r(p) - 1]\,\mathrm{d}p \tag{33}$$

perustuvan muodon[11]

$$EG_{NS} = \{\int_0^1 W(p)\,[r(p) - 1]^\eta\,\mathrm{d}p\}^{1/\eta}, \quad \eta \neq 0, \tag{34}$$

[10] Tämä tarkoittaa sitä, että Gini-kerroin ei järjestä jakaumia samoin kuin keskimääräinen hyöty (1), jos $U(y)$ on derivoituva.

[11] Itse asiassa esitetty muoto on vielä yleisempi, mutta tämä muoto tuo paremmin esiin yleistyksen idean.

jossa $W(p)$ on painofunktio, $r(p)$ skaalattu tulo ja η parametri. Gini-kertoimen tapauksessa $W(p) = 2p$ ja $\eta = 1$ (ks. yhtälö (25)).

Tässä työssä tarkastellaan yhtälöön (24) perustuvaa Gini-kertoimen yleistystä, jonka esittivät Kakwani (1980b), Donaldson ja Weymark (1980) sekä Yitzhaki (1983). Laajennettu Gini-kerroin määritellään nyt yhtälönä

$$EG(y;\nu) = 1 - \nu(\nu - 1)\int_0^1 L(p)(1-p)^{\nu-2}\mathrm{d}p, \tag{35}$$

jossa y ilmaisee tutkitun jakauman ja ν on parametri, joka kuvaa eriarvoisuuden kaihtamisen voimakkuutta. Parametri saa arvoja väliltä $\nu \in [1, \infty]$.[12] Parametrin arvolla $\nu = 2$ laajennettu Gini palautuu tavalliseksi Gini-kertoimeksi. Täten Gini-kerroin on erikoistapaus näin määritellyn laajennetun Gini-kertoimen indeksiperheestä. Parametrin arvoilla $\nu \in [1, 2)$ integroinnissa painotetaan korkeiden tulojen osaa jakaumasta ja arvoilla $\nu \in (2, \infty]$ matalien tulojen osaa. Tätä tarkastellaan myöhemmin lisää.

Yhtälön diskretisointi vaatii Lorenz-käyrän segmenttien integroinnin

$$EG(y;\nu) = 1 - \nu(\nu - 1)\sum_{j=1}^{n}\int_{p_{j-1}}^{p_j} L(p)(1-p)^{\nu-2}\mathrm{d}p, \tag{36}$$

jossa $p_0 = 0$ ja $p_m = 1$. Koska havaittu Lorenz-käyrä on havaintojen välillä suora, voidaan integrointi suorittaa helposti. Lyhyen laskun jälkeen (ks. Liite A) päädytään yhtälöön (Chotikapanich ja Griffiths 2001)

$$EG(y;\nu) = 1 + \sum_{j=1}^{n} r_j\left[(1-p_j)^{\nu} - (1-p_{j-1})^{\nu}\right], \tag{37}$$

jossa $r_j = y_j/\overline{y}$. Laajennettu Gini-kerroin voidaan määritellä myös käyttämällä yhtälön (31) mukaista kovarianssitulkintaa yhtälöllä (Lerman ja Yitzhaki 1995)

$$EG(y;\nu) = -\nu\,\mathrm{cov}(r, (1-p)^{\nu-1}), \tag{38}$$

jossa siis r on skaalattu tulo ja p kotitalouksien kertymä $F(y)$.

Laajennettu Gini-kerroin sallii siis erilaiset painotukset pieni- ja suurituloisille. Siten sen avulla voidaan tarkastella jakaumaa monipuolisemmin kuin pelkällä Gini-kertoimella. Näin määritellyn laajennetun Gini-kertoimen hyvinvointiteoreettisia ominaisuuksia ovat käsitelleet esimerkiksi Kakwani (1980a, 73–79) ja Lambert (2001, 125–126).

3.2.3 Atkinsonin indeksi

Atkinson (1970) esitti kuuluisassa artikkelissaan Lorenz-dominanssin (jakso 2.3.1) lisäksi ehdotuksen uudeksi eriarvoisuusmitaksi. Mitta perustuu yhteiskunnalle saman hyvinvoinnin tason tuottavaan tasaisesti jaettuun tuloon y^*, jonka määrittelee yhtälö

$$U(y^*) = \int_{y_{min}}^{y_{max}} U(y)\,dF(y). \tag{39}$$

[12] Kakwani (1980) esitti yhtälön parametrilla on $k = \nu - 1$, jolloin parametri väli on $k \in [0, \infty]$. Yitzhakin parametrointi on kuitenkin jäänyt käytännöksi.

Koska hyötyfunktio U on oletettu konkaaviksi, pätee relaatio $y^* < \overline{y}$. Tämän vuoksi on luonnollista määritellä eriarvoisuusmitta yhtälön

$$I(y) = 1 - \frac{y^*}{\overline{y}} \tag{40}$$

mukaisesti, jolloin tulon tasajako tuottaa indeksin arvon nolla ja eriarvoisuuden kasvu vähentää indeksin arvoa. Ratkaisemalla y^* yhtälöstä (39) saadaan indeksille diskreetissä muodossa yhtälö

$$I(y) = 1 - \frac{1}{\overline{y}}\, U^{-1}(\sum_{j=1}^{n} w_j U(y_j)). \tag{41}$$

On ilmeistä, että indeksin käyttämiseksi täytyy tehdä oletus hyötyfunktion U muodosta.

Atkinson (1970) käytti konkaavia hyötyfunktiota, jolla on vakioinen eriarvoisuuden kaihtamisen aste ϵ (constant inequality-aversion), eli hyötyfunktio on muotoa

$$U(y) = A + B\frac{y^{1-\epsilon}}{1-\epsilon}, \text{ kun } \epsilon \neq 1 \tag{42}$$

ja

$$U(y) = \log(y), \text{ kun } \epsilon = 1. \tag{43}$$

Jotta U olisi konkaavi, parametrille täytyy päteä $\epsilon > 0$. Sijoittamalla yllä olevat hyötyfunktiot yhtälöön (41) saadaan Atkinsonin indeksi esitettyä operationaalisessa muodossaan

$$I(y;\epsilon) = 1 - \overline{m}(r; 1-\epsilon), \tag{44}$$

jossa $\overline{m}$ on momenttikeskiarvo, joka on määritelty muodossa

$$\overline{m}(x;\tau) = \begin{cases} \left[\sum_{j=1}^{n} w_j x_j^{\tau}\right]^{1/\tau}, & \tau \neq 0 \\ \exp\left[\sum_{j=1}^{n} w_j \log(x_j)\right], & \tau = 0. \end{cases} \tag{45}$$

Momenttikeskiarvon määritelmän mukaisesti $\epsilon = 0$ vastaa aritmeettista keskiarvoa ja siten parametrin lähestyessä nollaa, lähestyy indeksin arvo nollaa. Parametrin arvo $\epsilon = 1$ vastaa geometrista ja $\epsilon = 2$ harmonista keskiarvoa. Koska parametrin ϵ kasvaessa τ pienenee, pienenee myös momenttikeskiarvon arvo. Siten Atkinsonin indeksin arvo kasvaa, kun ϵ kasvaa.

Atkinsonin indeksissä siis kuvataan tulojakauma $\mathbf{y}$ jakaumaksi $\mathbf{u}$ funktiolla U siten, että $u_i = U(y_i)$. Saadusta hyötyjakaumasta lasketaan keskiarvo, joka kuvataan takaisin tuloavaruuteen käänteiskuvauksella U^{-1}. Saatu tulo on siis tulo, joka tasaisesti jaettuna tuottaisi yhteiskunnalle saman hyvinvoinnin kuin mitä nykyinen jakauma $\mathbf{y}$ tuottaa. Tämä tulo y^* suhteutetaan nykyjakauman keskiarvoon, jotta saadaan suhteellinen eriarvoisuusindeksi.

Indeksin käyttö pakottaa olettamaan muodon eettisen tarkkailijan hyvinvointifunktiolle. Tämä ei kuitenkaan välttämättä ole huono ominaisuus. Kaikkien mittojen taustalla on jokin painorakenne, joka vastaa hyvinvointiteoreettisia hyvinvointifunktioita

(Sen 1973, 43). Tämä siitä huolimatta, että mitan käyttäjä ei tätä eksplisiittisesti ilmaise tai tiedosta. Atkinsonin indeksin etuna voidaankin pitää sitä, että painofunktio on eksplisiittisesti esitetty ja että siihen voidaan vaikuttaa parametrilla, jolla on selkeä tulkinta. Parametria ϵ muuttamalla tutkija voi tarkastella tulojakaumaa erilaisin painotuksin. Siten Atkinsonin indeksillä voidaan tutkia jakaumaa erilaisin painotuksin, kuten oli myös laajennetun Gini-kertoimen tapauksessa mahdollista. Itse asiassa mittojen välillä on tiettyjä yhteyksiä, joita on käsitellyt Yitzhaki (1983). Sen sijaan on totta, että käytetty painofunktio tuskin on yhteiskunnassa vallitseva hyvinvointifunktio, eikä myöskään ole helppo tehtävä selvittää sitä, millä parametrin arvolla painofunktio kuvaisi tätä parhaiten. Sen sijaan voidaan tehdä tarkastelut monilla parametrin arvoilla ja tehdä johtopäätökset moniin eri painotuksiin perustuen.

Indeksin käytettävyyttä rajaa se, että momenttikeskiarvo ei ole määritelty nollahavainnoille, kun $\epsilon \geq 1$. Tämä voi johtaa ongelmiin, tutkittaessa tuloeroja vuositason aineistolla. Lisäksi parametrin ϵ saadessa suuria arvoja, lähestyy hyvinvointifunktio rawlsilaista maximin-hyvinvointikäsitettä. Otosaineiston tapauksessa tämä ei ole mieluisa ominaisuus, sillä pienin havainto saa tuloksissa nopeasti hyvin suuren painon.

3.2.4 Yleistetty entropia

Entropia on monella alalla esiintyvä informaation määrän, epäjärjestyksen ja satunnaisuuden mitta. Informaatioteorian ja fysiikan käyttämä muoto on

$$H = -\sum_{j=1}^{n} p_j \log(p_j), \tag{46}$$

jossa p_j on tapahtuman j suhteellinen frekvenssi. Tilastotieteen hajontalukuna käyttämä entropia on määritelty puolestaan yhtälöllä (Vartia 1989, 99)

$$H' = -\sum_{j=1}^{n} p_j \log_2(p_j), \tag{47}$$

eli logaritmin kantaluku on e:n sijasta 2. Entropia saa maksimaalisen arvonsa, kun kaikki tapahtumat ovat yhtä todennäköisiä eli kun $p_j = 1/n$, kaikilla $j = 1, 2, 3, \ldots, n$, mikä vastaa täydellistä epäjärjestystä.

Informaatioteoriassa entropialla kuvataan havaintojoukon informaatiosisältöä laskemalla painotettu keskiarvo painojen funktiosta $h(p_j)$ (Theil 1967, jakso 2.1). Funktio h on sellainen, että se saa suuria arvoja, kun p_j on pieni. Näin funktio antaa suuren arvon niille havainnoille, jotka ovat epätodennäköisiä. Valinta $h = -\log$ perustuu siihen, että näin riippumattomien havaintojen yhteistodennäköisyys $p_i p_j$ voidaan hajottaa muotoon $h(p_i p_j) = h(p_i) + h(p_j)$ (Cowell 1995, 48). Fysiikassa puolestaan logaritmi johtaa siihen, että entropia on ns. ekstensiivinen suure ja että tasapainossa kaikki makrotilan muodostavat erilaiset mikrotilat ovat yhtä todennäköisiä (esim. Arponen 1994), kuten pitääkin olla.

Theil (1967) sovelsi ensimmäisenä informaatioteorian entropiaa eriarvoisuuden mittaamiseen. Hän esitti, että tapahtuman todennäköisyys p_j olisi korvattava havainnon tulo-osuudella $y_j/(N\overline{y})$. Indeksiksi hän esitti maksimaalisen entropian ja havaitun entropian erotusta. Näin määriteltynä päädytään indeksiin (ks. Liite A)

$$T = \sum_{j=1}^{n} w_j \, r_j \, \log(r_j), \tag{48}$$

jossa $r_j = y_j/\overline{y}$, $\overline{y} = \sum w_k y_k$, $w_j = f_j/N$ ja $N = \sum f_k$.

Theilin indeksi on erikoistapaus ns. yleistetyn entropian -indeksiperheestä (esim. Cowell ja Kuga 1981). Jos luovutaan edellä vaadituista informaatioteoreettisista (ja fysikaalisista) ominaisuuksista, voidaan $-\log$-funktio korvata funktiolla

$$h(x;\alpha) = \frac{1 - x^{\alpha-1}}{\alpha(\alpha-1)N^{1-\alpha}}, \tag{49}$$

jossa $\alpha \neq 0$ ja $\alpha \neq 1$. Parametrin arvolla $\alpha = 1$ funktiolla h on raja-arvo $h = -\log$ (Liite A), mutta arvolla $\alpha = 0$ sillä ei ole raja-arvoa.[13] Laskettaessa nyt Theilin ehdottamalla tavalla eriarvoisuusindeksi, käyttäen entropiassa funktion $h = -\log$ sijasta yhtälön (49) mukaista funktiota, päädytään (ks. Liite A) indeksiin $GE(\alpha)$, jota kutsutaan yleistetyksi entropiaksi[14]

$$GE(\alpha) = \frac{1}{\alpha(\alpha-1)} \left[\sum_{j=1}^{n} w_j \, r_j^{\alpha} - 1 \right]. \tag{50}$$

Tutkimalla raja-arvot $\alpha = 0$ ja $\alpha = 1$ havaitaan (ks. Liite A), että

$$GE(0) = -\sum_{j=1}^{n} w_j \, \log(r_j), \tag{51}$$

eli keskimääräinen log-poikkeama ja

$$GE(1) = \sum_{j=1}^{n} w_j \, r_j \, \log(r_j), \tag{52}$$

eli Theilin entropia-mitta T. Näiden raja-arvotarkastelujen lisäksi on syytä huomata, että

$$GE(2) = \frac{1}{2} CV^2, \tag{53}$$

jossa CV on jakauman variaatiokerroin.

Edellisessä jaksossa esitetty Atkinsonin eriarvoisuusmitta voidaan esittää yleistetyn entropian muunnoksena. Voidaan nimittäin osoittaa (ks. Liite A), että pätee yhtälö

$$I(\epsilon) = \begin{cases} 1 - [1 - \epsilon(1-\epsilon)GE(1-\epsilon)]^{1/(1-\epsilon)}, & \epsilon \neq 1 \\ 1 - \exp\left[-GE(0)\right], & \epsilon = 1. \end{cases} \tag{54}$$

[13] Lienee ilmeistä, että funktiolle h olisi voitu käyttää luonnollisempaakin yleistystä, esim. $h(x) = (1 - x^a)/a$, mutta yhtälön (49) mukainen yleistys on osoittautunut ominaisuuksiltaan mielenkiintoisimmaksi.

[14] Momenttikeskiarvoa käyttäen $GE(y;\alpha) = [\alpha(\alpha-1)]^{-1}[\overline{m}\,(r;\alpha)^{\alpha} - 1]$.

Yleistetty entropia on siis järjestysekvivalentti Atkinsonin indeksin kanssa, kun $\alpha = 1 - \epsilon$, sillä mittojen välillä on aidosti kasvava kuvaus. Koska $\epsilon > 0$, yleistetty entropia järjestää jakaumat samoin kuin Atkinsonin indeksi, kun $\alpha < 1$. Parametrin arvoilla $\alpha \geq 1$ GE-indeksillä voidaan tutkia jakaumaa sellaisin painotuksin, joihin Atkinsonin indeksi ei pysty. Yleistetyn entropian voidaan siis nähdä olevan Atkinsonin indeksin yleistys.

GE-indeksiperhe perustuu, kuten Atkinsonin indeksikin, momenttikeskiarvon käsitteeseen ja siten momenttikeskiarvo määrää sen, millä parametrin arvoilla mitta on määritelty nollahavaintojen tapauksessa. Yleistetty entropia on määritelty nollahavaintojen tapauksessa vain, kun $\alpha > 0$. Huomaa, että Theilin entropiaindeksi eli $GE(1)$ on sen sijaan määritelty, vaikka joukossa olisi nolla-havaintoja, sillä $\lim_{x \to 0} x \log(x) = 0$ (ks. Liite A). Yleistettyä entropiaa voidaan käyttää myös kun aineistossa on negatiivisia havaintoja. Tällöin parametrin on oltava positiivinen kokonaisluku.

3.3 Painotuksesta

Kuten edellisestä luvusta havaittiin, on soveltajalla suuri joukko eriarvoisuusmittoja, josta valita. Toisin kuin Lorenz-dominansseihin perustuvassa päättelyssä, tutkijan on nyt tiukennettava oletuksia, joiden vallitessa hänen tuloksensa pätevät. Tutkijan on siis otettava kantaa arvoihin, joiden mukaan järjestäminen tehdään. Tämä tapahtuu valitsemalla eriarvoisuusmittaperhe sekä tämän parametrin arvo. Jotta valinta voitaisiin perustellusti tehdä, tutkijan on tunnettava eriarvoisuusmittojen jakaumasta esiin tuomat ominaisuudet ja painotukset. Niihin tutustutaan tässä jaksossa.

3.3.1 Mittojen arvot, kun jakaumat ovat samat

Ensimmäinen kysymys, joka herää mittojen arvoja vertailtaessa, on: millaisia arvoja eri mitat saavat, kun ne lasketaan samasta jakaumasta ja miten arvot muuttuvat, kun jakaumaa muutetaan? Kysymykseen on näin aseteltuna mahdoton vastata, sillä tulos riippuu täysin tutkittavasta jakaumasta ja siitä, millainen muutos jakaumaan tehdään. Tämän vuoksi havainnollistetaan mittojen ominaisuuksia yksinkertaisimmalla mahdollisella, mutta silti mielenkiintoisella, jakaumalla. Tutkitaan jakaumaa, joka koostuu kahdesta osajoukosta, joiden kunkin jäsenet saavat saman tulon. Osajoukkojen osuudet jakaumasta ovat p_1 ja $1 - p_1$. Jakauman määrittävät siis parit (p_1, y_1) ja $(1 - p_1, y_2)$. Jotta eriarvoisuus vastaisi jakauman hyvinvointia, on tulojen keskiarvojen oltava jakaumilla samat eli pätee $p_1 y_1 + (1 - p_1) y_2 = \overline{y} =$ vakio. Jatkossa eri jakaumat esitetään nelikkoina $(p_1, r_1; p_2, r_2)$, jossa jo tunnetusti $r_i = y_i / \overline{y}$ ja siten $p_1 r_1 + p_2 r_2 = 1$.

Kuten edellisessä luvussa esitettiin, jakauman eriarvoisuutta kuvataan tyypillisesti Lorenz-käyrällä. Käyrän muoto voidaan yksikäsitteisesti esittää osuuden p_1 ja parametrin a avulla, kun $a = p_1 - L(p_1)$ eli Lorenz-käyrän taitepisteen pystysuora etäisyys tasajakosuorasta. Koska $L(p_1) = p_1 y_1 / \overline{y}$, tuloille y_1 ja y_2 saadaan yksikäsitteiset arvot

$$y_1 = \overline{y}\,(1 - \frac{a}{p_1}) \quad \text{ja} \quad y_2 = \overline{y}\,(1 + \frac{a}{1 - p_1}). \tag{55}$$

Sijoittamalla saadut tulot eriarvoisuusmittojen lausekkeisiin (37), (44) ja (50) saadaan mitoille lausekkeet

$$EG(\nu) = \frac{a}{p_1}(1 - (1 - p_1)^{\nu-1}), \quad \nu > 1, \tag{56}$$

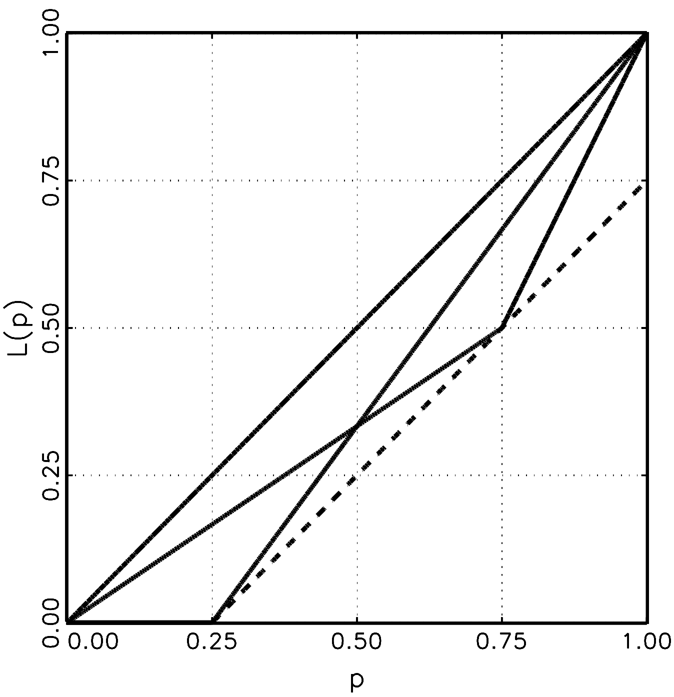
$$I(\epsilon) = 1 - \left[p_1 r_1^{1-\epsilon} + (1 - p_1) r_2^{1-\epsilon}\right]^{\frac{1}{1-\epsilon}}, \quad \epsilon > 0, \epsilon \neq 1 \tag{57}$$

ja

$$GE(\alpha) = \frac{1}{\alpha(1-\alpha)}\left[1 - p_1 r_1^{\alpha} - (1 - p_1) r_2^{\alpha}\right], \quad \alpha \neq 0, 1, \tag{58}$$

joissa r_i on skaalattu tulo $y_i/\overline{y}$, joiden arvot ovat suoraan yhtälöstä (55) $r_1 = 1 - a/p_1$ sekä $r_2 = 1 + a/(1 - p_1)$. Vastaavanlaiset esitykset saadaan myös parametrin arvoille, jotka johtavat rajankäynteihin. Yhtälöistä havaitaan, että eriarvoisuusmitat ovat yksikäsitteisesti määrättävissä parametrien a ja p_1 avulla.

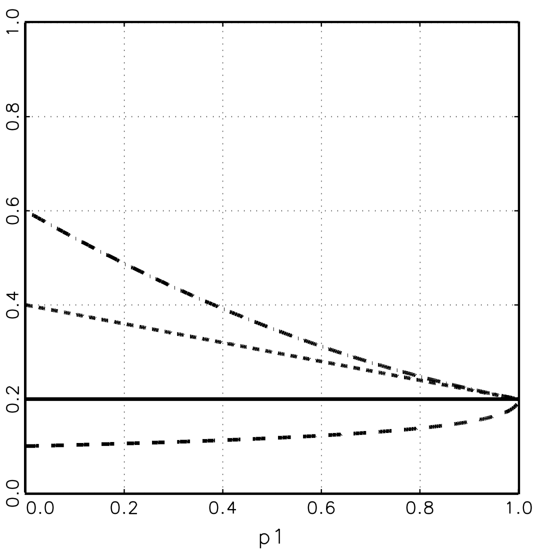
Edellisessä luvussa todettiin, että Gini-kerroin ei järjestä jakaumia samalla tavalla minkään derivoituvan hyvinvointi- ja hyötyfunktion kanssa. Tämä kielii siitä, että Gini-kertoimen määräämässä luokittelussa voi olla joitain ei-toivottuja piirteitä. Syy ongelmiin havainnollistuu, kun tutkitaan yhtälöä (56) parametrin arvolla $\nu = 2$. Gini-kerroin saa arvon a kaikilla p_1:n arvoilla, koska Lorenz-ala on kolmio, jonka leveys on yksi ja korkeus a. Esimerkiksi arvolla $a = 1/4$ Gini-kertoimen arvo on 1/4 jakaumille (1/4, 0; 3/4, 4/3) sekä (3/4, 2/3; 1/4, 2). Jakaumia vastaavat Lorenz-käyrät on esitetty kuvassa 1. Gini-kertoimen arvo voidaan siis tulkita prosenttiosuudeksi populaatiosta, jotka

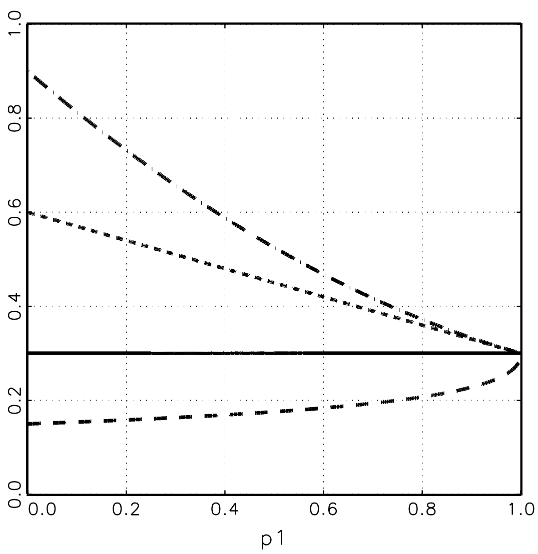


Kuvio 1 Jakaumien $(p_1, r_1; p_2, r_2) = (1/4,\ 0;\ 3/4,\ 4/3)$ sekä $(p_1', r_1'; p_2', r_2') = (3/4,\ 2/3;\ 1/4,\ 2)$ Lorenz-käyrät. Katkoviivalla on osoitettu niiden taitepisteiden ura, joille Gini-kerroin saa arvon 1/4. Siten myös esimerkkijakaumien Gini-kertoimet saavat arvon 1/4.

kahden ryhmän tapauksessa saavat nollatulon.[15] Näyttäisi siis siltä, että Gini-kerroin ei eriarvoisuusmitaksi tuo tarpeeksi voimakkaasti esiin suuren nollatulojen osuuden aiheuttamaa heikkoa hyvinvointia.

Koska Gini-kerroin saa kaikilla p_1 arvokseen parametrin a, on Gini-kerroin luonnollinen mittapuu muiden mittojen käyttäytymiselle. Kuvissa 2, 3 ja 4 on esitetty laajennetun Ginin, Atkinsonin ja yleistetyn entropian mittojen arvot eri parametrin a arvoilla p_1:n funktiona, kun tutkitut jakaumat koostuvat kahdesta osajoukosta kuvan 1 mukaisesti. Vertailtujen jakaumien $EG(2) = a$. Nyt siis p_1 on ainoa muuttuja, joka määrää Lorenz-käyrän muodon. Kuvasta 2 havaitaan, että laajennettu Gini parametrin arvoilla $\nu < 2$





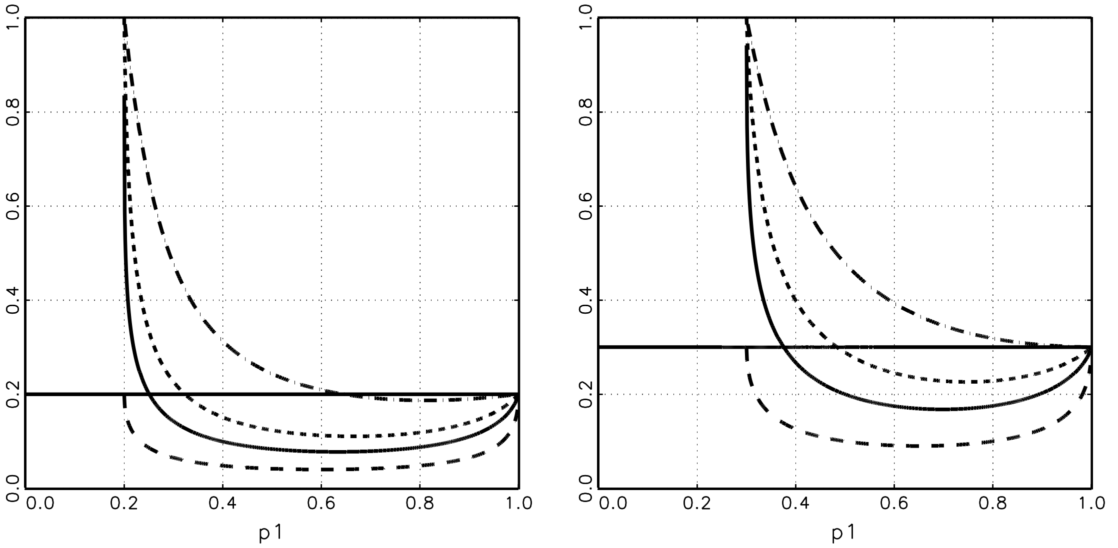
Kuvio 2 Laajennetun Gini-kertoimen arvot p_1:n funktiona, parametrin arvoilla $\nu = 1,5$ (katkoviiva), $\nu = 2$ (yhtenäinen viiva) $\nu = 3$ (lyhyt katkoviiva) ja $\nu = 4$ (katko-pisteviiva). Vasemmassa kuvassa $a = 0,2$, oikeassa $a = 0,3$.

pitää parhaana[16] jakaumia, joissa on suuria negatiivisia tuloja, eikä lainkaan erittäin suurituloisia. Parametrin arvoilla $\nu > 2$ tilanne on päinvastainen: optimaalinen jakauma on sellainen, jossa infinitesimaalista joukkoa lukuun ottamatta kaikki saavat samaa alle keskiarvon olevaa tuloa $r_1 = 1 - a$ ja jäljelle jäänyt pieni ryhmä saa rajattomasti kasvavat tulot. Siten parametrin arvolla $\nu > 2$ etenkin pienet tulot ovat eriarvoisuuden lähde ja arvolla $\nu < 2$ suuret.

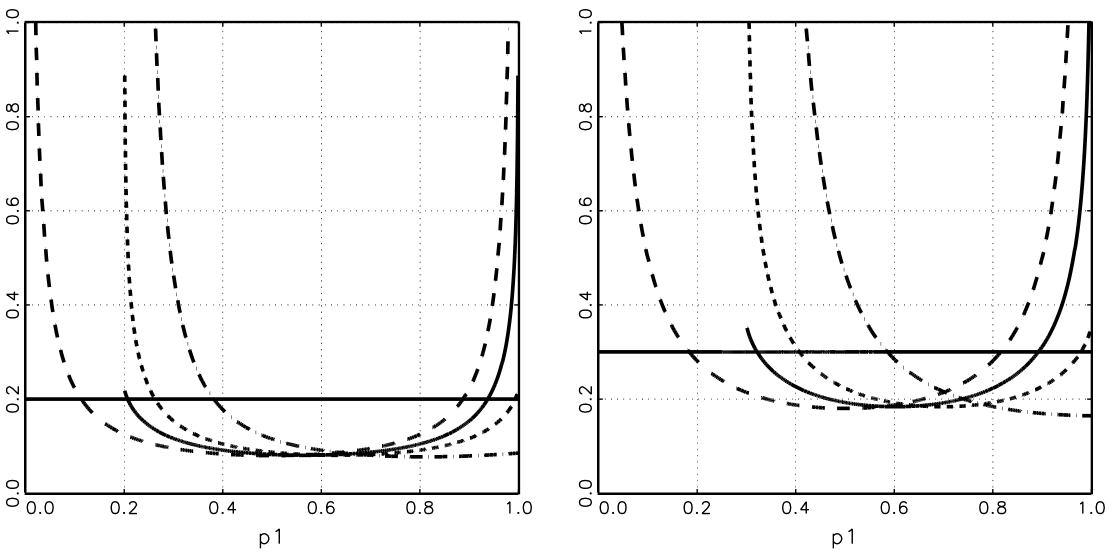
Atkinsonin indeksi järjestää jakaumat pieniä tuloja kaihtaen, kun $\epsilon > 1$, mutta järjestys ei ole muuten yhtä suoraviivainen kuin laajennetun Gini-kertoimen tapauksessa (kuva 3). Suurilla parametrin ϵ arvoilla indeksi näyttää painottavan tuloja samoin kuin laajennettu Gini-kerroin parametrin arvoilla $\nu > 2$, kun Gini-kertoimen arvo on tarpeeksi suuri. Gini-kertoimen kasvaessa, eli eriarvoisuuden määrän kasvaessa, yhä suurempi osa Atkinsonin indekseistä painottaa tuloja samoin kuin $EG(\nu > 2)$. Pienillä parametrin ϵ arvoilla Atkinsonin indeksi järjestää jakaumat siten, että parhaalla jakaumalla tulonjako on melko tasainen: jakaumassa ei ole hyvin pieniä, eikä hyvin suuria tuloja.

15 Tämän tulkinnan esittävät myös Nygård ja Sandström (1981, 244).

16 Mitan mielessä paras jakauma on sellainen, jossa mitta saa pienimmän mahdollisen arvon.



Kuvio 3 Atkinsonin indeksin arvot p_1:n funktiona, parametrin arvoilla $\epsilon = 0,5$ (katkoviiva), $\epsilon = 1$ (yhtenäinen viiva) $\epsilon = 1,5$ (lyhyt katkoviiva) ja $\epsilon = 3,5$ (katko-pisteviiva). Vasemmassa kuvassa $a = 0,2$, oikeassa $a = 0,3$. Vaakasuora yhtenäinen viiva osoittaa Gini-kertoimen arvon.



Kuvio 4 Yleistetyn entropian arvot p_1:n funktiona, parametrin arvoilla $\alpha = 2$ (katkoviiva), $\alpha = 1$ (yhtenäinen viiva) $\alpha = 0$ (lyhyt katkoviiva) ja $\alpha = -2,5$ (katko-pisteviiva). Vasemmassa kuvassa $a = 0,2$, oikeassa $a = 0,3$. Vaakasuora yhtenäinen viiva osoittaa Gini-kertoimen arvon.

Atkinsonin indeksillä on miellyttävä ominaisuus, että rajalla $p_1 = 1$ se saa arvon a. Rajalla $p_1 \to a$ pätee $I(\epsilon \geq 1) = 1$ ja $I(\epsilon < 1) \in (0,1)$.

Yleistetty entropiamitta $GE(\alpha)$ järjestää jakaumat parametrien arvoillaan $\alpha < 1$ kuten Atkinsonin indeksi $I(1-\alpha)$. Muulloin ($\alpha \geq 1$) yleistetty entropia järjestää jakaumat siten, että mitta saa suuria arvoja, kun havaintojen joukossa on keskiarvosta suuresti poikkeavia arvoja eli $|r_i| >> 1$. Nollahavainnot eivät tätä ehtoa täytä, joten mitta ei pidä nollatuloja erikoisen pieninä, sallien niiden esiintymisen.

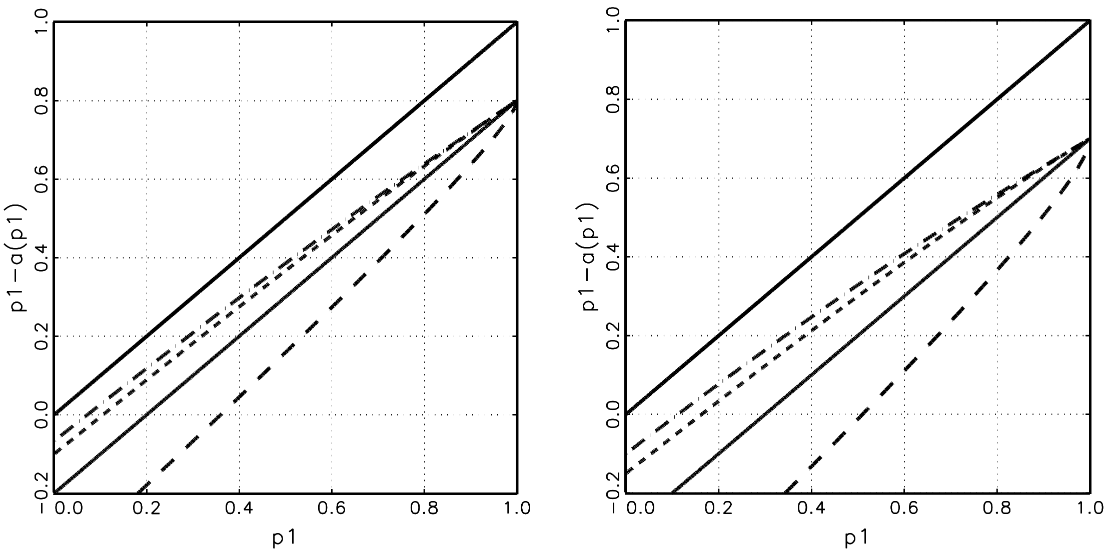
Tutkitaan kuvaa 2 vielä konkreettisen esimerkin avulla. Oletetaan, että tulojakauma on ensin muotoa (p_1,r_1;p_2,r_2) = (0,2, 0; 0,8, 1,25), mutta muuttuu siitä tilaan (p'_1,r'_1;p'_2,r'_2) = (0,4, 0,5; 0,6, 4/3). Gini-kerroin on kummassakin tapauksessa 0,2, joten sen avulla ei muutosta edes havaita. Parametrin arvolla $\nu = 4$ muutos on 0,488 $\to$ 0,392 eli eriarvoisuus väheni. Parametrin arvolla $\nu = 1,5$ muutos on 0,106 $\to$ 0,113, eli eriarvoisuus kasvoi. Mitä eriarvoisuudelle todella kävi, on normatiivinen päätös, mutta mittojen muutoksista voi päätellä sen, että hyvin pienituloisten osuus on laskenut ja suurituloisten kasvanut. Jos Gini-kerroinkin olisi muuttunut, olisivat muilla parametreilla havaittavat muutokset antaneet silloinkin tietoa muutosten rakenteesta. Tämä pätee tietenkin myös muihin eriarvoisuusmittoihin.

3.3.2 Saman mitan arvon tuottavat jakaumat

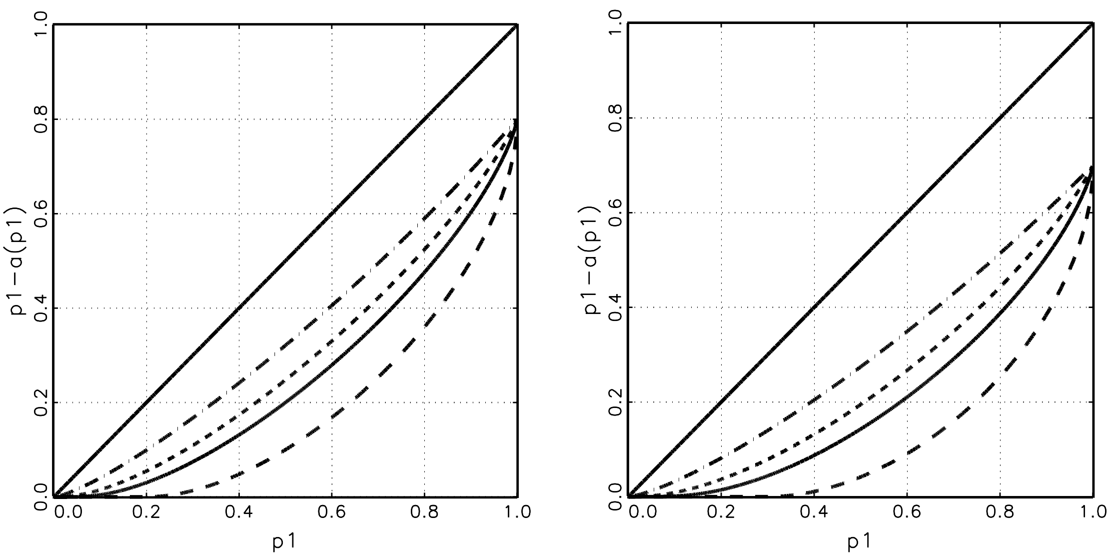
Edellä tutkittiin sitä, millaisia eri mittojen arvoja samasta jakaumasta saadaan. Nyt mielenkiinnon kohteena on se, millaisille jakaumille (sama) mitta tuottaa saman arvon. Kysymykseen etsitään vastausta edellisessä jaksossa esitellyllä konstruktiolla, jossa jakauma muodostuu kahdesta osajoukosta. Kun edellä pidettiin vakiona Lorenz-käyrän taitepisteen pystysuoraa etäisyyttä tasajakosuorasta a, niin nyt etsitään a:n arvo p_1:n funktiona, kun mitan arvo pidetään vakiona I_0. Jakaumaa voidaan jälleen siis kuvata yksikäsitteisesti parin (p_1, a) avulla. Käytännössä jakaumaa on havainnollisempi kuvata parin $(p_1, p_1 - a)$ avulla, joka on Lorenz-käyrän taitepiste.

Parametri a saadaan ratkaistua p_1:n funktiona yhtälöistä (56) – (58) asettamalla mitan arvoksi vakio I_0. Kuviin 5, 6 ja 7 on esitetty Lorenz-käyrän taitepisteiden $(p_1, p_1 - a(p_1))$ urat eri mittaperheille. Vasemman puoleisissa kuvissa $I_0 = 0,2$ ja $I_0 = 0,3$. Kuvassa 5 on esitetty laajennetun Gini-kertoimen tapaus. Kuvasta nähdään, kuinka parametrin ν kasvu painottaa integroinnissa enemmän pienten tulojen osaa Lorenz-alasta. Siten parametrin kasvaessa Lorenz-käyrän taitepiste siirtyy lähemmäksi tasajakosuoraa pienillä kotitalouskertymillä p_1 kuin suurilla. Mitä pienempiä tuloja pienituloisten ryhmässä saadaan, sitä pienempi Lorenz-ala tarvitaan saman mitan arvon tuottamiseksi, kun $\nu > 2$. Gini-kertoimen tapauksessa tunnetusti Lorenz-alan koko ei vaihtele. Jos puolestaan $\nu < 2$, niin tilanne on päinvastainen verrattuna tilanteeseen $\nu > 2$. Tällöin suuret tulot ovat pääasiallinen eriarvoisuuden lähde.

Kuvassa 6 on esitetty tilanne Atkinsonin indeksin tapauksessa. Kun parametri ϵ saa suuren arvon ja Lorenz-ala on suuri, havaitaan samantapainen käyttäytyminen kuin laajennetun Gini-kertoimen tapauksessa parametrin ν ollessa suuri: pienet tulot ovat eriarvoisuuden tärkein lähde. Lorenz-käyrän taitepisteen ura on kuitenkin konveksi, kun Ginin tapauksessa se oli konkaavi. Pienemmillä parametrin arvoilla pienet ja suu-



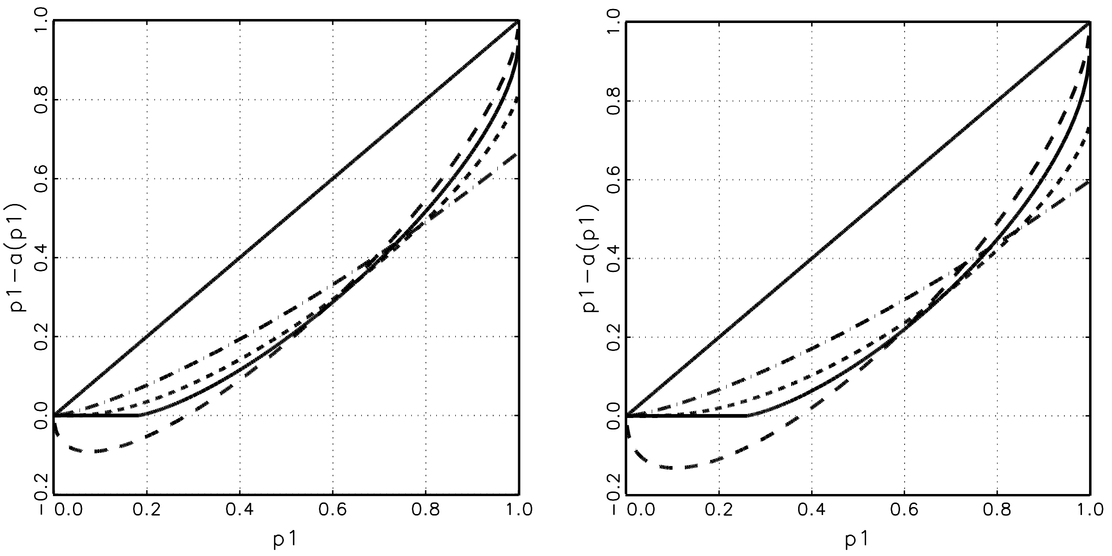
Kuvio 5 Kahden osajoukon populaation Lorenz-käyrän taitepisteen ura, joka tuottaa saman laajennetun Gini-kertoimen arvon I_0 parametrin arvoilla $\nu = 1,5$ (katkoviiva), $\nu = 2$ (yhtenäinen viiva) $\nu = 3$ (lyhyt katkoviiva) ja $\nu = 4$ (katkopisteviiva). Vasemmassa kuvassa $I_0 = 0,2$, oikeassa $I_0 = 0,3$.



Kuvio 6 Kahden osajoukon populaation Lorenz-käyrän taitepisteen ura, joka tuottaa saman yleistetyn entropian arvon I_0 parametrin arvoilla $\epsilon = 0,5$ (katkoviiva), $\epsilon = 1$ (yhtenäinen viiva) $\epsilon = 1,5$ (lyhyt katkoviiva) ja $\epsilon = 3,5$ (katkopisteviiva). Vasemmassa kuvassa $I_0 = 0,2$, oikeassa $I_0 = 0,3$.

ret tulot ovat tasasuhtaiset eriarvoisuuden lähteet. Atkinsonin indeksin tapauksessa pieneksi tuloksi lasketaan skaalattu tulo, joka on positiivinen, mutta ykköstä huomattavasti pienempi.

Kuvassa 7 on esitetty yleistetyn entropian tapaus. Kuten edellä on mainittu, vastaa yleistetty entropia parametrin arvoilla $\alpha < 1$ ordinaalisesti Atkinsonin indeksiä para-



Kuvio 7 Kahden osajoukon populaation Lorenz-käyrän taitepisteen ura, joka tuottaa saman Atkinsonin indeksin arvon I_0 parametrin arvoilla $\alpha = 2$ (katkoviiva), $\alpha = 1$ (yhtenäinen viiva) $\alpha = 0$ (lyhyt katkoviiva) ja $\alpha = -2,5$ (katkopisteviiva). Vasemmassa kuvassa $I_0 = 0,2$, oikeassa $I_0 = 0,3$.

metrilla $\epsilon = 1 - \alpha$, joka havaitaan myös kuvista 6 ja 7. Sen sijaan muilla parametrin α arvoilla saadaan jälleen erilainen painotus. Rajatapauksessa $\alpha = 1$ ja sitä suuremmilla parametrin arvoilla mitan mukaan eriarvoisuutta aiheuttavat niin suuret kuin pienetkin tulot. Nyt mittapuuna on kuitenkin skaalatun tulon absoluuttinen poikkeaminen ykkösestä, eli nollatulo ei ole mitan mielestä vielä järin pieni tulo. Tämä näkyy kuvassa hyvin parametrin $\alpha = 2$ tapauksessa, jossa keskiarvotulon suuruusluokkaa olevien negatiivisten tulojen jälkeen alkaa Lorenz-ala pienentyä.

3.3.3 Herkkyys poikkeaville havainnoille

Edellä tehtyjen jakaumaa koskevien tarkastelujen lisäksi on hyvä tarkastella sitä, kuinka suuren painon eri suuruiset tulot mitassa saavat. Tämä kertoo luonnollisella tavalla siitä, millaisia tuloja mitta painottaa ja onko painotukset sellaisia, että poikkeavan suuruiset havainnot voivat yksin vaikuttaa oleellisesti mitan arvoon. Kuten mittojen määritelmistä havaittiin, perustuvat Atkinsonin indeksi sekä yleistetty entropia momenttikeskiarvoon. Tällöin tulojen painotus tehdään potenssiin korottamalla. Tällöin hyvin suuri tai pieni arvo voi potenssiin korotuksen jälkeen saada suhteettoman suuren arvon, mikä turmelee mitan antamat tulokset. Laajennetussa Gini-kertoimessa painotus tapahtuu eri tavoin. Siinä tuloa ei manipuloida, vaan sen sijaintia jakaumassa. Tämän etuna on se, että kertymäfunktio saa arvoja nollan ja yhden väliltä, joten tilanne on edeltä käsin rajoitettu.

Tutkitaan ensin momenttikeskiarvoon perustuvia mittoja. Näissä mitoissa skaalattu tulo on muodossa r_i^τ. Siten mitta on herkkä poikkeavan pienille $r << 1$ tuloille, kun τ on iso negatiivinen luku. Toisaalta poikkeavan suuri havainto $r >> 1$ on ongelmallinen, kun τ on suuri luku. Tämä on mahdollista vain yleistetyn entropian tapauksessa. Tarkastellaan ongelmaa nyt lisäämällä tunnettuun jakaumaan uusi havaintoyksikkö, jonka skaalattu tulo on r_0. Jos mitan arvo muuttuu lisäyksen johdosta suuresti, voidaan arvella mitan olevan herkkä poikkeaville havainnoille. Herkkyyden määräävät suuren muutoksen aiheuttava r_0:n suuruus sekä muutoksen rajuus. Nyt tehtävissä tarkasteluissa kaikkien havaintojen paino on $1/n$.

Lisättäessä uusi havainto $r_0 = y_0/\overline{y}^{vanha}$ alkuperäiseen jakaumaan, muuttuu luonnollisesti tulojen keskiarvo. Siten vanhat skaalatut tulot muuttuvat yhtälön

$$r_j^{uusi} = r_j^{vanha} A(n, r_0) \tag{59}$$

mukaisesti, jossa funktio A on keskiarvojen muuntofunktio

$$A(n, r_0) = \frac{\overline{y}^{vanha}}{\overline{y}^{uusi}} = \frac{n+1}{n+r_0}. \tag{60}$$

Asettamalla mitan lähtöarvoksi I_0, voidaan lyhyen laskun jälkeen tulon r_0 lisäyksen jälkeiset mitan arvot esittää muodossa

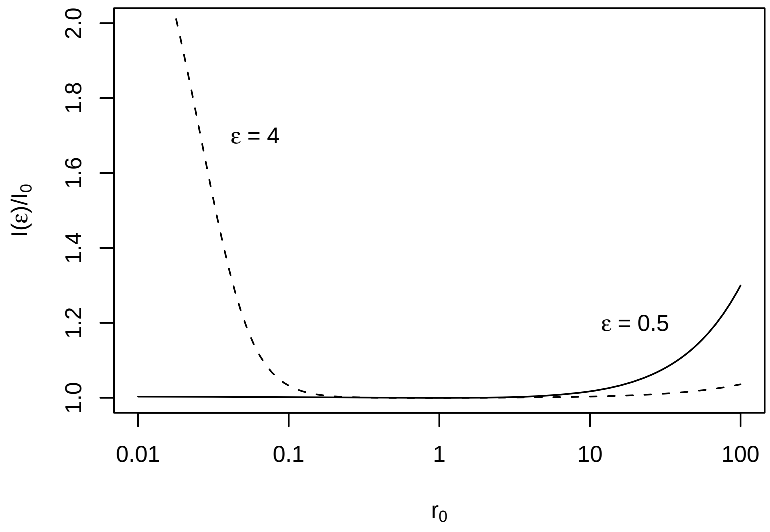
$$\frac{I(y^{uusi};\epsilon)}{I_0} = \frac{1}{I_0}\left\{1-(1-I_0)A(n,r_0)\left[\frac{n+\left(\frac{r_0}{1-I_0}\right)^{1-\epsilon}}{n+1}\right]^{\frac{1}{1-\epsilon}}\right\} \tag{61}$$

ja

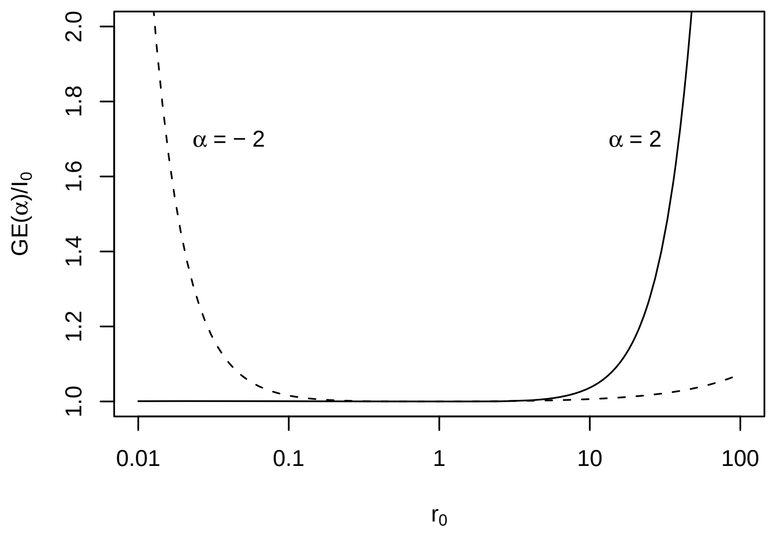
$$\begin{aligned}\frac{GE(y^{uusi};\alpha)}{I_0} &= \frac{1}{\alpha(\alpha-1)I_0}\left[n(A(n,r_0)^{\alpha}-1)+A(n,r_0)^{\alpha}r_0^{\alpha}-1\right]+\\ &\quad \frac{n}{n+1}A(n,r_0)^{\alpha}.\end{aligned} \tag{62}$$

Havaitaan, että muutoksen suuruus riippuu havaintojen määrästä n, alkuperäisestä mitan arvosta I_0 ja tietenkin lisätystä skaalatusta tulosta r_0. Kuvissa 8 ja 9 esitetty yleistetyn entropian ja Atkinsonin indeksin arvojen muutokset kahdella parametrin arvolla, lisätyn skaalatun tulon funktiona, kun havaintojen lukumäärä $n = 5000$. Kuvista havaitaan, että Atkinsonin indeksi suurilla parametrin arvoilla on hyvin herkkä pienille havainnoille. Pienillä parametrin arvoilla ei ole vastaavaa ongelmaa, vaan suurilla r_0:n arvoilla havaittu kasvu voitaneen pitää hyväksyttävän suuruisena. Yleistetyn entropian tapauksessa (kuva 9) poikkeamat käyvät nopeasti huolestuttavan suuriksi. Siten otosaineiston kanssa on rajoituttava pienelle parametrivälille.

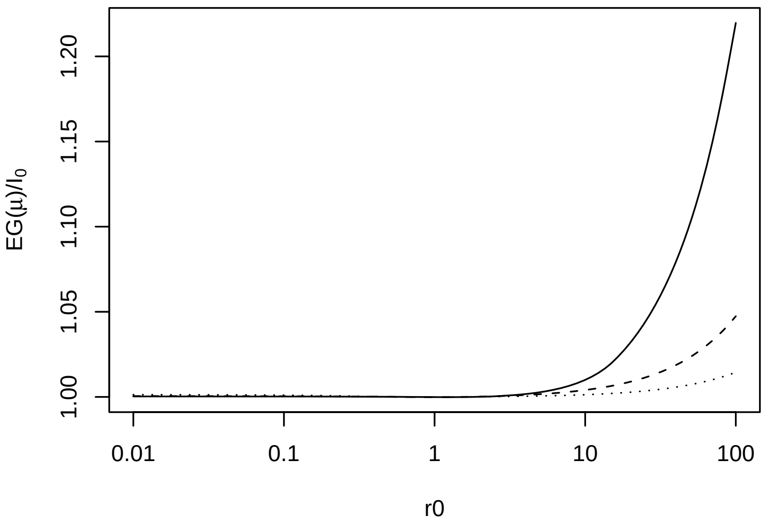
Laajennetun Gini-kertoimen tapauksessa ei yhtä suora lähestyminen ole mahdollista, koska tarkasteluja varten on tehtävä oletuksia alkuperäisestä jakaumasta. Pelkkä eriarvoisuusmitan arvo ei nyt nimittäin riitä, vaan tarvitaan lisäksi tieto tulojen sijainnista jakaumassa. Siksi tässä tarkastelu tehdään käyttämällä vuoden 2001 kulutustutkimuksen käytettävissä olevien tulojen jakaumaa, johon sijoitetaan edellä kuvattu havaintoyksikkö, jonka skaalattu tulo on r_0 ja paino $1/(n+1)$. Kuvasta 10 havaitaan, että hyvin suuret havainnot muuttavat mitan arvoa vain hieman, joten voidaan sanoa, ettei suurista poikkeavista havainnoista ole suurta huolta edes pienillä parametrin arvoilla. Kuvasta nähdään myös se, että laajennettu Gini-kerroin ei pidä pieniäkään positiivisia tuloja niin pieninä, että ne suuresti vaikuttaisivat tuloksiin. On syytä muistaa, että lähtöjakaumassa on mukana nollahavaintojakin, joten yksittäisten pienten positiivisten havaintojen ei pitäisikään vaikuttaa merkittävästi mitan arvoon.



Kuvio 8 Skaalatun tulon r_0 lisäämisen jälkeisen Atkinsonin indeksin suhde alkuperäiseen parametrin arvoilla $\epsilon = 4$ (katkoviiva) ja $\epsilon = 0.5$ (yhtenäinen viiva). I_0 on valittu siten, että se vastaa Suomessa havaittuja arvoja: $I_0 = 0,05$, kun $\epsilon = 0,5$ ja $I_0 = 0,35$, kun $\epsilon = 4$.



Kuvio 9 Skaalatun tulon r_0 lisäämisen jälkeisen yleistetyn entropian suhde alkuperäiseen parametrin arvoilla $\alpha = -2$ (katkoviiva) ja $\alpha = 2$ (yhtenäinen viiva). I_0 on valittu siten, että se vastaa Suomessa havaittuja arvoja: $I_0 = 0,2$ kummassakin tapauksessa.



Kuvio 10 Skaalaun tulon r_0 lisäämisen jälkeisen laajennetun Ginin suhde alkuperäiseen parametrin arvoilla $\nu = 1.001$ (yhtenäinen viiva), $\nu = 2$ (katkoviiva) ja $\nu = 10$ (pisteviiva). Lähtöjakaumana on vuoden 2001 käytettävissä olevien tulojen jakauma, jossa negatiiviset tulot on muutettu nolliksi. I_0:n arvot ovat: $I_0(\nu = 1.001) = 0,0006$, $I_0(\nu = 2) = 0,27$ ja $I_0(\nu = 10) = 0,55$.

3.4 Gini-kertoimen harhasta

Aiemmin mainittiin, että otosaineistoa käytettäessä Lorenz-ordinaattojen yhdistäminen suoralla pitää sisällään implisiittisen oletuksen siitä, että havaintojen väliin jäävällä populaation osalla on kaikilla sama tulo. Näin ei asia tietenkään ole, vaan näin laskettu Gini-kerroin antaa alaspäin harhaisen kuvan tilanteesta: havaittu eriarvoisuus on liian vähäistä.

Ongelma on luonnollisesti sitä suurempi, mitä vähemmän havaintoja on käytössä. Erityisesti käytettäessä desiiliaineistoa, on näin tehty virhe huomattavan suuri. Käytettäessä laajaa otosaineistoa ongelma ei ole niin akuutti, mutta sen suuruus on syytä arvioida. Cowell (1995, 108–116) käsittelee ongelmaa luokitellun aineiston tapauksessa, kun luokkien ylä- ja alarajat sekä luokan keskimääräinen arvo tunnetaan. Havaitaan nimittäin, että luokassa tasaisesti jakautuneen tulon oletus johtaa eriarvoisuuden alarajaan ja sopiva jako kahteen eri suureen tuloon johtaa mitan ylärajaan. Menetelmä ei kuitenkaan toimi mikroaineiston tapauksessa, jossa tunnetaan vain luokkien keskimääräiset tulot ja siten edellä mainittua sopivaa jakoa kahteen eri suureen tuloon ei voida tehdä.

Erittäin karkea arvio eriarvoisuuden ylärajalle saadaan, kun korvataan suorilla yhdistetty Lorenz-käyrä askelfunktiovastineella. Tämä menetelmä antaa aivan liian ison arvion Gini-kertoimelle, mutta suurten havaintomäärien tapauksessa sekin tarkentuu oikeaksi. Tarkastellaan siis nyt, kuinka suuren Gini-kertoimen arvon menetelmä antaa.

Jos arvo poikkeaa vain vähän alarajasta, voidaan alaraja-arvion käyttöä pitää riittävänä.

Implisiittisen luokkien sisäisen tasajako-oletuksen ja askelfunktio-approksimaation välinen ero on

$$\delta G = \sum_{j=1}^{n} w_j^2 r_j, \tag{63}$$

joka on Lorenz-ordinaattoja yhdistävien suorien alle jäävien suorakulmaisten kolmioiden ala kerrottuna kahdella. Harha ei siis ole läheskään näin suurta, mutta näin laskemalla saadaan sille ehdoton yläraja. Jos oletetaan, että painot ovat suuruudeltaan $w_i \approx 1/n$ kaikilla i, niin harhan ylärajaksi saadaan

$$\delta G \approx 1/n, \tag{64}$$

koska standarditulon keskiarvo on yksi. Siten karkean askelfunktio-approksimaation erotus nyt käytettyyn Gini-kertoimen lausekkeeseen on kääntäen verrannollinen havaintomäärään. Koska havaintoja on käytössä yli tuhat, on askelfunktio-Ginin virhe suuruusluokkaa yksi tuhannesosa. Näin voidaan päätellä, että koska käytetyn laskutavan harha on tätä huomattavasti pienempi, ei harhaa tarvitse ottaa tarkasteluissa huomioon.

Deltas (2003) on Monte Carlo -simulaatioin osoittanut, että pienillä otoksilla Gini-kerroin on alaspäin harhainen ainakin tasa-, log-normaali- ja eksponentiaalijakauman tapauksissa. Tämä ongelma koskee pieniä, lähinnä kymmenien tai sitä pienemmän havaintomäärän, aineistoja, eikä ole siten mikrotason aineistojen tapauksessa merkittävä. Ongelma voi kuitenkin nousta esiin, tehtäessä tuloerotutkimuksia aineiston osaryhmille, joissa on vain vähän havaintoja.

3.5 Muita mittojen ominaisuuksia

Eriarvoisuusmitoille on määritelty monia muitakin ominaisuuksia, joita voidaan pitää hyvän mitan kriteereinä. Tässä niistä mainitaan kaksi. Ensimmäinen on ns. vahva siirtoperiaate (esim. Cowell 1995, 61–65), jonka mukaan tulonsiirron suurituloiselta pienituloiselle pitäisi tasoittaa eriarvoisuutta sitä enemmän mitä suurempi henkilöiden tulo-osuuksien ero on. Tällaisen ehdon toteuttaa yleistetty entropia. Esimerkiksi Gini-kertoimen tapauksessa eriarvoisuuden muutos riippuu kotitalouksien osuudesta, joka on siirtoon osallistuvien välissä, eikä siten toteuta ehtoa.

Usein on mielenkiintoista tutkia sitä, kuinka jakaumassa havaittu eriarvoisuus muodostuu populaation osaryhmien sisäisestä ja välisestä eriarvoisuudesta. Tällöin käytetyltä eriarvoisuusmitalta voidaan vaatia ns. additiivisen hajotettavuuden ominaisuus. Tällä tarkoitetaan sitä, että jos jakauman F eriarvoisuus on jakaumaa G vähäisempää, on tällöin myös jakaumassa $\delta F + (1-\delta)H$ vähemmän eriarvoisuutta kuin jakaumassa $\delta G + (1-\delta)H$, kun jakauma H on mielivaltainen ja $\delta \in (0,1)$. Jälleen yleistetty entropia ja sen kanssa ordinaalisesti ekvivalentit mitat toteuttavat ehdon, mutta esimerkiksi Gini-kerroin ei (Cowell 1995, 57–60). Toinen tapa jakaa jakauma osiinsa, on tarkastella tulokomponentteja, joista tutkittu tulokäsite summaamalla muodostuu. Tällöin puolestaan Gini-kertoimelle voidaan johtaa mielekäs additiivinen hajotelma tulokomponenttien Gini-kertoimista (Lerman ja Yitzhaki 1985).

4 Hajonnan estimoinnista

Kun aineistosta lasketaan tilastollinen tunnusluku, on saadun piste-estimaatin luotettavuus luonnollinen mielenkiinnon kohde. Tyypillinen lähtökohta luotettavuuden arvioimiselle on estimaatin keskihajonta, joka kuvaa otoksesta johtuvaa estimaatin tyypillistä vaihtelun määrää. Jotta edellisessä luvussa kuvattuja tunnuslukuja olisi mielekästä käyttää, on niille pyrittävä estimoimaan myös hajontaestimaatit. Niiden avulla voimme tarkastella sitä, onko luvuissa havaittavat erot satunnaisvaihtelua, vai onko tutkituissa jakaumissa tapahtunut tilastollisesti merkitseviä muutoksia. Tässä luvussa esitellään tutkittujen eriarvoisuusmittojen asymptoottisia jakaumaominaisuuksia sekä uudelleenotantaan perustuvia hajonnan estimointimenetelmiä.

4.1 Asymptoottiset jakaumat

Kun tutkittu tunnusluku on laskukaavaltaan riittävän yksinkertainen, voidaan tunnusluvun hajonnalle johtaa yleinen lauseke. Esimerkiksi keskiarvo on tällainen tunnusluku. Sen sijaan edellisessä luvussa käsitellyt eriarvoisuusmitat ovat laskukaavoiltaan liian monimutkaisia, jotta niille voitaisiin johtaa yleiset keskihajonnan lausekkeet. Mitat ovat kuitenkin riittävän säännöllisesti käyttäytyviä, jotta otoksen kasvaessa rajatta, päädytään yleisen raja-arvoväittämän määrittämään jakaumaan. Tämän asymptoottisen jakauman keskihajonnan lauseketta voidaan käyttää keskihajonnan mittana myös äärellisen otoksen tapauksessa, mutta tällöin kyseessä on approksimaatio, jonka riittävyyttä on vaikea arvioida.

4.1.1 Laajennettu Gini-kerroin

Barrett ja Pendakur (1995) ovat esittäneet laajennetun Gini-kertoimen asymptoottisen jakauman. Lähtökohtana on mitan esitysmuoto

$$EG(y;\nu) = 1 - \sum_{j=1}^{n} a_j \Phi_j, \tag{65}$$

jossa

$$a_i = n\left[(1-p_{i-1})^\nu - 2(1-p_i)^\nu + (1-p_{i+1})^\nu\right], \quad i \in \{1,2,\ldots,n-1\} \tag{66}$$

ja

$$a_n = n^{1-\nu}. \tag{67}$$

Yhtälössä p_i on kotitalouksien kertymä ja Φ_i on Lorenz-ordinaatta (ks. jakso 3.2.1). Koska laajennettu Gini-kerroin voidaan näin esittää Lorenz-ordinaattojen lineaarikombinaationa, voidaan Gini-kertoimen asymptoottisissa jakaumatarkasteluissa käyttää hyväksi Lorenz-käyrien tunnettuja tuloksia (Beach ja Davidson 1983).

Jotta tulos voitaisiin esittää ymmärrettävästi, on määriteltävä lisää merkintöjä. Jaetaan aineisto $K+1$:een kvantiiliin, ja merkitään abskissoja p_i, jossa $i \in \{1,2,\ldots,K+1\}$

($p_{K+1} = 1$). Vastaava ordinaatta on Φ_i, joka voidaan esittää muodossa

$$\Phi_i = \sum_{j=1}^{i} w_j \frac{y_j}{\mu} = p_i \frac{\mu_i}{\mu} = p_i R_i, \tag{68}$$

jossa ensimmäinen yhtäsuuruus seuraa yhtälöstä (9), kun μ on tulojen odotusarvo. Toinen yhtäsuuruus seuraa määritelmästä $\mu_i = \sum_{k=1}^{i} w_k y_k / p_i$. Viimeinen yhtä suuruus määrittelee keskiarvojen suhteen $R_i = \mu_i / \mu$ (vrt. $r_i = y_i / \overline{y}$). Ordinaattaa Φ_i vastaavaa tuloa merkitään y_{p_i}, joka saadaan yhtälöstä $F(y_{p_i}) = p_i$, jossa F on tulojen kertymäfunktio. Yhtälössä (68) esiintyvä μ_i on tuloon y_{p_i} rajoittuvien tulojen odotusarvo. Tämän lisäksi tarvitaan vielä ehdollinen varianssi σ_i^2, joka on tuloon y_{p_i} rajoittuvien tulojen varianssi. Siten $\mu = \mu_{K+1}$ ja $\sigma^2 = \sigma_{K+1}^2$.

Edellä esitettyjen merkintöjen avulla laajennetun Gini-kertoimen asymptoottinen jakaumatulos voidaan riittävien säännöllisyysehtojen voimassa ollessa esittää muodossa:

$$\sqrt{n}\left[\widehat{EG}(\nu) - EG(\nu)\right] \xrightarrow{D} \mathrm{N}(0, V), \tag{69}$$

jossa asymptoottinen varianssi

$$V = \sum_{i=1}^{K} \sum_{j=1}^{K} a_i a_j v_{ij}, \tag{70}$$

jossa a_i on määritelty yhtälön (66) mukaisesti,

$$v_{ij} = \frac{1}{\mu^2}(\omega_{ij} + p_i R_i p_j R_j \sigma^2 - p_i R_i \omega_{j,K+1} - p_j R_j \omega_{i,K+1}), \tag{71}$$

ja

$$\omega_{ij} = p_i \left[\sigma_i^2 + (1 - p_j)(y_{p_j} - \mu_i)(y_{p_j} - \mu_j) + (y_{p_i} - \mu_i)(\mu_j - \mu_i)\right]. \tag{72}$$

Termi v_{ij} voidaan yhtälön (68) mukaisesti kirjoittaa muodossa

$$v_{ij} = \frac{1}{\mu^2}(\omega_{ij} + \Phi_i \Phi_j \sigma^2 - \Phi_i \omega_{j,K+1} - \Phi_j \omega_{i,K+1}) \tag{73}$$

ja sen tulkinta on Lorenz-ordinaattojen Φ_i ja Φ_j asymptoottinen kovarianssi. Yhtälö (73) on voimassa, kun $i \le j \in \{1, 2, \dots, K\}$ ja yhtälö (72), kun $i \le j \in \{1, 2, \dots, K+1\}$. Lausekkeissa esiintyvät teoreettiset suureet voidaan estimoida tarkentuvasti tyypillisillä otossuureillaan, joten myös otosvarianssi voidaan estimoida tarkentuvasti.

4.1.2 Atkinsonin indeksi ja yleistetty entropia

Myös Atkinsonin indeksille sekä yleistetylle entropialle on johdettu asymptoottiset jakaumatulokset (Thistle 1990).[17] Tulokset perustuvat origomomenttien asymptoottisiin

[17] Myös Cowell (1989) on esittänyt tavan estimoida yleistetyn entropian varianssia.

tuloksiin ja mittojen momenttiesityksiin. Muuttujan y α:s momentti on määritelty yhtälöllä

$$\mu(\alpha) = EY^\alpha = \int_0^\infty y^\alpha \mathrm{d}F(y), \tag{74}$$

kun integraali on olemassa. Nyt y:n on oltava positiivinen, sillä α:n ei tarvitse olla positiivinen kokonaisluku. Luonnollinen estimaattori momentille on otosmomentti

$$m(\alpha) = \sum_{j=1}^{n} w_j y_j^\alpha. \tag{75}$$

Merkitään lisäksi $\mathrm{Var}\,[\sqrt{n}m(\alpha)] = \sigma^2(\alpha)$ ja $\mathrm{Cov}\,[\sqrt{n}m(\alpha), \sqrt{n}m(\alpha')] = \gamma(\alpha, \alpha')$. Eli $\mu(1) = \mu$ ja $\sigma^2(1) = \sigma^2$ ovat muuttujan y odotusarvo ja varianssi.

Tällöin momenttien ollessa olemassa pätee: (a) havaintojen kasvaessa rajatta $m(\alpha) \to \mu(\alpha)$ todennäköisyydellä 1, (b) $\sigma^2(\alpha) = \mu(2\alpha) - \mu(\alpha)^2$, (c) $\sqrt{n}\,[m(\alpha) - \mu(\alpha)]$ ja $\sqrt{n}\,[m(\alpha') - \mu(\alpha')]$ ovat asymptoottisesti yhdessä normaalisesti jakautuneita odotusarvoilla nolla, kohdan (b) mukaisilla variansseilla sekä kovarianssilla $\gamma(\alpha, \alpha') = \mu(\alpha + \alpha') - \mu(\alpha)\mu(\alpha')$ (Thistle 1990).

Atkinsonin indeksistä käytetään momenttikeskiarvo-esitystä

$$I(\theta) = 1 - \frac{\mu(\theta)^{1/\theta}}{\mu(1)}, \tag{76}$$

jossa siis $\theta = 1 - \epsilon$ ja jota estimoidaan otossuureella

$$\widehat{I}(\theta) = 1 - \frac{m(\theta)^{1/\theta}}{m(1)}. \tag{77}$$

Kun $\theta \neq 0$, pätee $\widehat{I}(\theta) \xrightarrow{p} I(\theta)$ ja $\sqrt{n}\left[\widehat{I}(\theta) - I(\theta)\right] \xrightarrow{D} N(0, v_I(\theta))$, jossa

$$v_I(\theta) = \left(\frac{1 - I(\theta)}{\theta\mu(\theta)}\right)^2 \left(\sigma^2(\theta) - 2\frac{\theta\mu(\theta)}{\mu}\gamma(\theta, 1) + \frac{\theta\mu(\theta)}{\mu}\sigma^2\right) \tag{78}$$

(Thistle 1990). Atkinsonin indeksin estimaattori on siis tarkentuva. Varianssin lausekkeessa esiintyvät teoreettiset suureet voidaan estimoida luonnollisilla otossuureillaan.

Yleistetty entropia voidaan puolestaan kirjoittaa momenttien avulla muodossa

$$GE(\alpha) = k\left(\frac{\mu(\alpha)}{\mu^\alpha} - 1\right) \quad \alpha \neq 0, 1, \quad k = \alpha(\alpha - 1) \tag{79}$$

jolloin estimaattori on

$$\widehat{GE}(\alpha) = k\left(\frac{m(\alpha)}{m(1)^\alpha} - 1\right). \tag{80}$$

Tällöin pätee, että $\widehat{GE}(\alpha) \xrightarrow{p} GE(\alpha)$ ja $\sqrt{n}\left[\widehat{GE}(\alpha) - GE(\alpha)\right] \xrightarrow{D} N(0, v_G(\alpha))$, jossa

$$v_G(\alpha) = k^2\mu^{-2(\alpha+1)}\left[\mu^2\sigma^2(\alpha) - 2\alpha\mu\mu(\alpha)\gamma(\alpha, 1) + \alpha^2\mu(\alpha)^2\sigma^2\right] \tag{81}$$

(Thistle 1990). Myös tämä estimaattori on siis tarkentuva. Otosvarianssi saadaan jälleen korvaamalla teoreettiset suureet otosvastineillaan ja huomioimalla otoskoko n.

4.2 Uudelleenotanta

Edellä kuvatut jakaumatarkastelut pätevät suurten havaintomäärien rajalla ja johtavat siitä huolimatta, etenkin Gini-kertoimen tapauksessa, hankalasti laskettaviin lausekkeisiin. Tämän vuoksi voi olla tarpeen ottaa käyttöön laskennallisesti helpommat, ja tilastotieteelliseltä taustaltaan aivan toisenlaiset, ns. uudelleenotantamenetelmät.

Uudelleenotannassa käytössä olevasta otosaineistosta muodostetaan uusia aineistoja. Uuden aineiston muodostamistapa riippuu käytetystä menetelmästä, mutta menetelmästä riippumatta uudet aineistot kuvaavat alkuperäisessä otannassa mahdollisia lopputuloksia. Näin käyttöön saadaan mahdollisesti suurikin määrä aineistoja, joiden avulla voidaan estimoida mm. otantavirhettä.

Uudelleenotantaan perustuvat menetelmät ovat tyypillisesti ei-parametrisia ja laskentaintensiivisiä. Siten tehokkaiden tietokoneiden aikana niiden avulla saadaan helposti ratkaistua hankaliakin estimointitehtäviä turvautumatta jakaumaoletuksiin tai työlääseen ja monessa tilanteessa riittämättömään asymptotiikkaan. Tämän helppouden ja monikäyttöisyyden vuoksi menetelmille onkin annettu luonnetta kuvaavia nimiä kuten Jackknife (linkkuveitsi) ja Bootstrap (saapasremmi[18]).

4.2.1 Jackknife

Jacknife-menetelmässä uudelleenotanta tehdään siten, että aineistoon tehdään ositus osajoukkoihin, joissa on kussakin k havaintoa. Jokainen osajoukko poistetaan kerran aineistosta ja jäljelle jääneitä käytetään laskentaan. Näin saatuja n/k kappaletta aineistoja voidaan käyttää mm. varianssin estimointiin. Käytännössä parhaaksi ryhmän kooksi on havaittu $k = 1$, eli aineistoja saadaan otoskoon n verran. Tässä esityksessä tarkastellaan vain tätä tapausta. Hyvä yleisesitys menetelmästä on mm. Wolter (1965, 153–200), johon tämän jakson esitys myös osin perustuu.

Jackkife-menetelmä on alunperin esitetty estimaattoriksi, jolla saadaan vähennettyä tunnusluvun estimoinnissa esiintyvää harhaa (Quenouille, 1949 ja 1956). Tukey (1958) esitti, että menetelmällä voitaisiin ehkä laskea myös varianssiestimaatteja ja soveltaa testiteoriaa. Vuonna 1962 hän esitti julkaisemattomassa käsikirjoituksessa käsitteet pseudoarvo ja menetelmälle nimen Jackknife.

Tarkastellaan tunnuslukua t. Kun otoksesta poistetaan havainto i, tunnuslukua t vastaava ns. pseudoarvo τ_i määritellään yhtälöllä

$$\tau_i = nt - (n-1)t_i, \tag{82}$$

jossa t_i on tunnusluku t laskettuna ilman havaintoa i. Tunnusluvun t jackknife-estimaatin määritelmä on

$$\tau = \frac{1}{n}\sum_{j=1}^{n}\tau_j, \tag{83}$$

[18] Nimi pohjautuu paroni von Münchausenin seikkailuihin, joissa paroni pelastautui nostamalla itsensä saappaanremmistään.

eli pseudoarvojen samapainoinen keskiarvo. Jackknife-estimaatti voidaan kirjoittaa myös alkuperäisen tunnusluvun avulla sijoittamalla yhtälö (82) yhtälöön (83)

$$\tau = nt - (n-1)\bar{t}, \tag{84}$$

jossa $\bar{t} = \sum t_j/n$ eli samapainoinen keskiarvo. Jos pseudoarvot ovat korreloimattomia, voidaan jackknife-estimaatin varianssia estimoida yhtälöllä

$$\mathrm{var}_{jack}\,\tau = \frac{1}{n(n-1)}\sum_{j=1}^{n}(\tau_j - \tau)^2. \tag{85}$$

Tunnusluvun t jackknife-estimaatin varianssiestimaattia voidaan käyttää tunnusluvun t varianssin estimaattorina, joten sijoitettessa edelliseen yhtälöön yhtälöt (82) ja (84) saadaan tulos

$$\mathrm{var}_{jack}\,t = \frac{n-1}{n}\sum_{j=1}^{n}(t_j - \bar{t})^2. \tag{86}$$

Yllä esitetyt estimaattorit ovat asymptoottisesti hyviä monissa erilaisissa sovelluksissa (Miller 1974), mutta epälineaaristen tunnuslukujen tapauksessa teoreettiset perustelut jackknifen toimivuudelle ovat vähäiset (Wolter 1985, 171–172). Jackknife-varianssiestimaattorin soveltuvuutta myös t:n varianssin estimointiin on perustellut teoreettisesti Brilinger (1966), mutta käytännössä menetelmän käyttökelpoisuus on perusteltu etenkin Monte Carlo-simulaatioin (esim. Efron 1981) sekä läheisellä relaatiolla seuraavassa jaksossa esiteltävän bootstrap-menetelmän kanssa (esim. Efron ja Gong 1983).

Otoskoon kasvaessa myös laskujen määrä kasvaa. Tämän vuoksi voi olla tarpeen kehittää laskentatehokkaita algoritmeja kunkin tunnusluvun jackknife-estimaattorien laskentaan. Tämä rajoittaa menetelmän käytettävyyttä. Kyseessä on kuitenkin vain ohjelmointiongelma, eikä sitä siksi käsitellä tässä esityksessä.

Ongelmistaan huolimatta menetelmä antaa usein riittävän hyviä, simulaatioiden perusteella konservatiivisia, varianssiestimaatteja hyvinkin hankalasti laskettavien tunnuslukujen tapauksessa. Menetelmä ei toimi estimoitaessa tunnuslukuja, jotka eivät ole tarpeeksi sileitä funktioita kuten esim. mediaanin tapauksessa (Efron ja Tibshirani 1993, 148). Lisäksi on syytä huomata, että varsinaista jackknife-estimaaattia τ ei esimerkiksi eriarvoisuusindeksejä laskettaessa tarvita. Mielenkiinnon kohteena on vain menetelmän suoma varianssiestimaatti.

4.2.2 Bootstrap

Bootstrap (Efron 1979) on jackknife-menetelmän sukuinen uudelleenotantaan perustuva estimointimenetelmä, jota voidaan käyttää mm. hajonnan ja luottamusvälien estimointiin.[19] Menetelmän lähtökohtana on havaittu empiirinen jakauma $\hat{F}$, jossa jokaiselle havainnolle on annettu paino $1/n$. Tämä empiirinen jakauma on taustalla olevan

[19] Tässä esityksessä keskitytään vain hajonnan estimointiin.

jakauman F estimaatti. Mielenkiinnon kohteena on tunnusluvun θ estimaattorin $\hat{\theta}$ hajonta eli

$$\sigma(F; n, \hat{\theta}) = \left(\int [\, \hat{\theta}(y) - E_F\, \hat{\theta}(y)\,]^2 dF(y) \right)^{1/2} \tag{87}$$

jossa n on otoskoko. Hajonnan bootstrap-estimaattori puolestaan on

$$\hat{\sigma} = \sigma(\hat{F}; n, \hat{\theta}). \tag{88}$$

Tyypillisesti hajonnoista jätetään tunnettuina pois estimaattori ja otoskoko eli $\hat{\sigma} = \sigma(\hat{F})$. (Efron ja Tibshirani 1986.)

Hajonnan $\sigma(F)$ lauseke on useissa tilanteissa vain hyvin hankalasti esitettävissä, mutta tästä huolimatta sen bootstrap-estimaattori voidaan suoraviivaisesti laskea käyttäen Monte Carlo -simulointia (MC). Simuloinnin ensimmäisessä vaiheessa generoidaan B kappaletta n-alkioisia bootstrap-aineistoja $\mathbf{y}_b$ käyttäen alkuperäiseen aineistoon satunnaisotantaa palauttaen. Toisessa vaiheessa lasketaan bootstrap-replikaatit, eli tunnusluvun estimaatit $\hat{\theta}_b$ kullekin bootstrap-aineistolle $\mathbf{y}_b$. Kolmannessa ja viimeisessä vaiheessa replikaateista lasketaan hajontaestimaatti

$$\hat{\sigma} = \left[\frac{1}{B-1} \sum_{b=1}^{B} \left(\hat{\theta}_b - \overline{\theta} \right)^2 \right]^{1/2}, \tag{89}$$

jossa $\overline{\theta} = 1/B \sum_{b=1}^{B} \hat{\theta}_b$. Bootstrap-aineistoja tarvitaan tyypillisesti 25–200. (Efron ja Tibshirani 1993, 47). Vähäisemmän uudelleenotanta-aineistojen määrän vuoksi bootstrap on suurten aineistojen tapauksessa jackknife-menetelmää laskentatehokkaampi ja siksi sitä tullaan käyttämään tässä tutkimuksessa.

Toisin kuin jackknife-menetelmässä, bootstrap-simulaatiossa valitaan satunnaisesti osa havainnoista mukaan replikaatioiden laskuun. Yksittäisten replikaatioiden välinen suuruusero on näin bootstrap-menetelmässä suurempi kuin jackknifessa, ja siten se tuottaa vähemmällä laskumäärällä tarkemman arvion hajonnalle. Lisäksi jackknife-menetelmässä voidaan valita vain n erilaista otosta, josta replikaatiot lasketaan, kun taas bootstrap-menetelmän simulaatioissa bootstrap-otokseen voi tulla hyvin kirjava joukko havaintoja. Siten jackknife-menetelmä on eräässä mielessä lineaarinen approksimaatio bootstrap-menetelmästä (Efron ja Tibshirani 1993, luku 20).

Lisäksi on syytä huomata, että bootstrap-menetelmän MC-simulaatio johtaa siihen, että keskihajonnan arvot vaihtelevat ajokertojen välillä. Siten bootstrap-otosten määrä ja estimaatin tarkkuus on valittava siten, ettei valitulla tarkkuudella estimaatissa havaita eroja eri ajokertojen välillä.

Bootstrap-menetelmän rajoitukset ovat yhteneviä jackknife-menetelmän kanssa. Tärkein rajoite on se, että jos havaittu jakauma ei ole hyvä estimaatti todellisesta jakaumasta tutkitulla alueella. Tällaisia ongelmakohtia ovat etenkin jakauman päät. Toisaalta ongelmia tuottavat myös ei-sileät tunnusluvut. Esimerkiksi jakaumassa olevien atomien määrään, bootstrap antaa estimaatiksi n (kun havaintoja on n), mikä ei ole järin tyydyttävää (Efron ja Tibshirani 1993, 81).

4.2.3 Ei-vakioiset painot ja uudelleenotanta

Käytettäessä otosaineistoja, joissa havainnoilla on erisuuruiset painot, tarvitsevat edellä kuvatut uudelleenotantamenetelmät lisätäsmennyksen. Nyt nimittäin painojen erilaisuus on otettava huomioon udelleenotannassa. Jackknife-menetelmässä painot huomioidaan tunnuslukujen laskemisessa ja kukin havainto jätetään pois kerran. Bootstrapmenetelmässä on ongelmaan periaatteessa kaksi ratkaisua.

1. Bootstrap-otannassa annetaan kullekin havainnolle poimintatodennäköisyydeksi $1/n$. Replikaatin arvoa laskettaessa käytetään havainnoilla olevia korotuskertoimia f_i.

2. Bootstrap-otannassa annetaan kullekin havainnolle poimintatodennäköisyydeksi havainnon paino w_i. Replikaatin laskussa havainnot ovat samapainoisia $1/n$.

Ensimmäisessä menetelmässä havainnot ovat otannassa samanarvoisia ja painot huomioidaan testisuureen laskussa. Toisessa menetelmässä painot huomioidaan heti otannassa. Verrattaessa menetelmien tuloksia havaittiin, että ensimmäinen menetelmä antaa suuremmat keskivirhe-estimaatit, joskaan erot eivät olleet suuria. Konservatiivisuusperiaatetta noudattaen käytetään tässä tutkimuksessa menetelmää 1.

5 Aineisto

Tutkimuksen aineiston muodostavat Tilastokeskuksen vuosien 1985, 1990, 1995, 1998 ja 2001 viisi kulutustutkimusta, joista vuoden 1995 tutkimus sisältää vuosien 1994, 1995 ja 1996 pieniotoksiset tutkimukset (ks. esim. Ahlqvist ja Pajunen 2000, 88). Kulutustutkimuksia on tehty noin viiden vuoden välein vuodesta 1966, mutta edellä mainitut viisi tutkimusta ovat sisällöltään toisiinsa verrattavissa. Tämän vuoksi tutkimuksessa tyydytään tarkastelemaan tuloerokehitystä vuodesta 1985 vuoteen 2001.

Kulutustutkimukset ovat otospohjaisia kotitaloustason aineistoja, joihin on kerätty tulo- sekä kulutustietoja. Sekä tulo- että kulutustiedot sisältävät myös laskennallisia eriä. Kulutukseksi lasketaan muun muassa asuntoedun tuoma asumispalvelu sekä itse tuotetut että kerätyt tuotteet. Hinnat tällaisille kulutuserille määrätään markkinahinnoista. Laskennallisia tuloja ovat puolestaan esimerkiksi omistusasunnosta muodostuneet laskennalliset tulot. Tiedot on kerätty kolmella menetelmällä: haastatteluilla, tilinpidolla sekä hallinnollisista vuositason rekistereistä poimimalla. Tulotiedot on kerätty rekistereistä ja haastatteluin. Kulutustiedot on kerätty kahden viikon tilinpitoon perustuvista menoista sekä lähinnä kestohyödykkeitä täydentävistä haastatteluista. Kulutusmenotiedot on luokiteltu kansainvälisen COICOP-kulutusnimikkeistön[20] mukaisesti. (Tilastokeskus 2001.)

5.1 Tulokäsitteet

Kotitaloudet saavat tuloja monista lähteistä. Lopullinen käytettävissä oleva tulo muodostuu useiden tulokäsitteiden summana. Taulukossa 1 on esitelty, miten kulutustutkimuksessa kotitalouden eri tulokäsitteet koostuvat. Tuotannontekijätulot muodostuvat markkinoilla määräytyvistä tuloista. Kun tuotannontekijätuloihin lisätään kotitalouksien saamat tulonsiirrot, saadaan bruttotulot. Kotitaloudet maksavat tulonsiirtoja (lähinnä välittömiä veroja). Nämä bruttotuloista vähentämällä saadaan käytettävissä olevat tulot. Tuloa, joka saadaan vielä vähentämällä välilliset verot, kutsutaan veropuhdistetuksi tuloksi. Tämän suureen avulla tutkimuksessa arvioidaan kokonaisverotaakan jakautumista.

Käytettävissä olevat tulot (nettotulot) kuvaavat tulokäsitteistä tarkimmin kotitalouden taloudellisen hyvinvoinnin määrää. Toinen vaihtoehto olisi tarkastella kulutusmenoja. Kulutusmenojen voidaan sanoa noudattavan paremmin kotitalouden pysyväistuloa, joka on kotitalouden hyvinvoinnin teoreettinen lähde. Toisaalta kulutusmenoja tarkastelemalla ei havaita käyttämättömän tulon osan eli säästämisen tuottamaa hyvinvointia. Tulo- ja kulutuseriarvoisuusongelmaa pohditaan tarkemmin seuraavassa luvussa. Tuloeroja tutkitaan usein myös muiden tulokäsitteiden tasolla. Tällöin ei kuitenkaan ole järkevää antaa laskelmille eriarvoisuustulkintaa, sillä tarkastelussa on vain osa kotitalouksien tulovirrasta.

[20] Classification of Individual Consumption by Purpose. Lisätietoja luokituksesta on julkaissut mm. Tilastokeskus (2002).

Taulukko 1 Tulokäsitteet ja niiden muodostaminen.

	palkkatulot
+	yrittäjätulot
=	työ/ansiotulot
+	omaisuustulot
=	tuotannontekijätulot
+	saadut tulonsiirrot
=	bruttotulot
−	maksetut tulonsiirrot
=	käytettävissä olevat tulot (nettotulot)
−	välilliset verot
=	veropuhdistettu tulo

5.2 Aineiston ongelmat

Tuloerotutkimukseen paremmin soveltuva aineisto olisi luonnollisesti Tilastokeskuksen tulonjakotilasto. Taloudellisen eriarvoisuuden kannalta on kuitenkin perusteltua tarkastella myös kotitalouksien kulutusmenojen jakaumaa. Lisäksi työssä tutkitaan välillisten verojen kohdentumista ja välillinen vero muodostuu kulutuksen perusteella. Tulojen ja kulutuksen havaintokohtaista yhteyttä ei ole muissa aineistoissa kuin kulutustutkimuksissa. Tutkimuksen aineistona on siis käytettävä kulutustutkimuksia. Aineistolla on rajoituksensa, mutta se on riittävä menetelmien testaamiseen sekä yleiskuvan saamiseen.

Kulutustutkimusten yhtenä ongelmana voidaan pitää sitä, että otantavaihtelut huomioidenkin poikkeavat sen eräät tulotiedot vastaavista tulonjakotilastojen tiedoista. Poikkeamat ovat suurimpia eräissä tulonsiirtojen sekä yrittäjä- ja omaisuustulojen luokissa. Vaikka koko aineiston tasolla ongelmat eivät ole suuria, niin taustamuuttujien luokissa tehtävät tarkastelut voivat olla epäluotettavia. Lisäksi kulutustutkimukset eivät sisällä mm. lähdeveron alaisia ja verovapaita korkotuloja eikä useimpia kotitalouksien välisiä tulonsiirtoja (Tilastokeskus 2001, 12).

Tämän lisäksi kulutushavainnot poikkeavat joissain hyödykeluokissa yleensä huomattavasti alaspäin kansantalouden tilinpidon kotitaloussektorin vastaavista arvoista. Tällaisia poikkeavia ryhmiä ovat etenkin alkoholijuomat, tupakka, rahoituspalvelut sekä ravintola- ja hotellipalvelut.[21] On myös hyödykeryhmiä, jotka ovat kulutustutkimuksissa suurempia kuin kansantalouden tilinpidossa. Nämä puutteet vaikuttavat luonnollisesti välillisten verojen vaikutusten tarkasteluun ja on siten muistettava ottaa huomioon tuloksia analysoitaessa (Tilastokeskus 2001, 27–30). Lisäksi hyvinvoinnin kannalta oleelliset julkiset hyvinvointipalvelut puuttuvat kulutustutkimuksen kulutustiedoista.

Oman ongelmansa muodostaa myös se jo aiemmin mainittu seikka, että aineiston tulot on kerätty yhden vuoden aikana. Koska etenkin yrittäjä- ja omaisuustulot voivat vaihdella huomattavastikin vuosien välillä, kasvattaa tämä satunnaisvaihtelu havaittua hajontaa. Jos tulojen lyhytkestoisen satunnaisvaihtelun määrä on muuttunut vuosien

[21] Esimerkiksi alkoholin osalta kulutustutkimuksen tietojen suhde kansantalouden tilinpidon tietoihin oli 46 % vuonna 2001 (Tilastokeskus 2001, 49).

saatossa, näkyy tämä tuloerojen muutoksina, vaikka hyvinvoinnissa ei olisikaan tapahtunut muutoksia. Periaatteessa tulonjakoon liittyvät eriarvoisuuspohdinnat pitäisi suorittaa elinkaaritulojen tai Friedmannin pysyväistulon perusteella. Suomen tapauksessa tutkijalla on käytössään kuitenkin vain yhden vuoden tulotiedot.[22]

Oman lisäongelmansa tuovat aineistossa olevat negatiiviset havainnot. Yrittäjä- ja omaisuustulon tapauksessa negatiiviset tulot ovat luonnollisia, koska tuloerät ovat nettomääräisiä. Kun tutkitaan palkka-, brutto- sekä käytettävissä olevia tuloja, negatiiviset havainnot ovat sekä käsitteellisesti (palkkatulot) että hyvinvointitarkastelujen kannalta (kotitalouden kokonaistulot vuoden ajalta negatiivisia) ongelmallisia. Lisäongelman muodostaa se tekninen murhe, että Atkinsonin indeksi ja yleistetty entropia eivät tyypillisesti ole määriteltyjä negatiivisilla tulon arvoilla. Tämän välttämiseksi em. negatiiviset tulot on muutettu tarkasteluissa nollatuloiksi. Osassa parametrialueestaan em. mittarit eivät ole määriteltyjä myöskään nollahavaintojen tapauksessa ja tällöin tyydytään tarkastelemaan positiivisten tulojen jakaumaa. On syytä huomata, että tämä rajaa suuresti mitan sovellettavuutta. Liitteeseen B on taulukoitu negatiivisten ja nollahavaintojen suhteelliset osuudet eri tulokäsitteillä eri kulutustutkimuksissa.

5.3 Eroavuudet tarpeissa

Käsiteltäessä kulutustutkimusten aineistoja on jakaumavertailua tehtäessä huomioitava se, että kotitaloudet ovat rakenteeltaan (esim. kooltaan ja ikärakenteeltaan) ja siten myös tarpeiltaan erilaisia. Kuusihenkisen talouden tarpeet ovat suuremmat kuin lapsettoman pariskunnan. Siten hyvinvointitarkasteluissa ei riitä tarkastella pelkkiä kotitalouksien tuloja, vaan on pyrittävä aidosti huomioimaan tulojen tuottama hyvinvointi. Tässä työssä hyvinvointivaikutukset huomioidaan käyttämällä ns. ekvivalenssiskaaloja e.[23]

5.3.1 Ekvivalenssiskaaloista

Kuten edellä mainittiin, kotitaloudet ovat erilaisia ja siten myös niiden tarpeet ovat erilaisia: eri määrä tuloa vaaditaan erilaisten kotitalouksien yhdenvertaiseen elintasoon. Jotta erilaisia kotitalouksia voitaisiin vertailla, pitäisi kotitaloudet pystyä kuvaamaan standardi-mittayksiköissä.

[22] Tulonjakotilastoissa perättäisten vuosien havaintoyksiköistä noin puolet on mukana myös seuraavana vuonna. Tämän ns. rotatoivan paneeliaineiston avulla voidaan tutkia sitä, kuinka paljon havaitut tuloerot muuttuvat, kun tutkitaankin kotitalouksille kahden vuoden aikana kertyneiden tulojen jakaumaa. Tämä keskiarvoistus kuvaa hieman paremmin pysyväistuloa, sillä se vaimentaa tulojen lyhytkestoista vaihtelua. Riihelä ja Sullström (2002, Taulukko 4) ovat tulonjakotilastoa käyttäen osoittaneet, että käytettävissä olevien tulojen tapauksessa kahden vuoden tietojen käyttö alentaa vain hieman saatuja Gini-kertoimen arvoja. Vaikuttaisi siis siltä, että vuositulojen käyttö ei aiheuta suuria ongelmia eriarvoisuustulkinnoille.

[23] Toinen tapa lähestyä ongelmaa on ns. perättäisten dominanssien kriteeri (Atkinson ja Bourguignon 1987). Menetelmä on luonnollinen lähestymistapa etenkin Lorenz-dominanssien tapauksessa, mutta tähän tutkimukseen ekvivalenssiskaalat sopivat paremmin.

Ekvivalenssiskaaloja käytettäessä oletetaan, että tämä standardointi voidaan suorittaa. Tyypillisesti standardimitaksi valitaan yksin elävä aikuinen ihminen eli ns. ekvivalenttiaikuinen. Ekvivalenssiskaala on funktio, joka kertoo, kuinka paljon kotitaloudessa on kulutustarvetta, mittayksikön ollessa yksin elävä aikuinen. Funktio on rakennettu siten, että se huomioi kotitalouden erilaisten jäsenten erilaiset kulutustarpeet sekä kotitalouden toimintaan liittyvät mittakaavaedut.

Ekvivalenssiskaala voidaan esittää muodollisesti funktiona

$$e_i = e(\mathbf{x}(i)), \tag{90}$$

joka siis kuvaa kotitalouden i tarpeet määrittävät ominaisuudet $\mathbf{x}(i)$ luvuksi e_i, joka on se yksin elävien aikuisten määrä, jolla on yhteensä samat tarpeet kuin kotitaloudella i. Siten standardoinnin jälkeen kotitaloudella ei ole tarpeiden suhteen muita ominaisuuksia kuin ekvivalenttiaikuisten lukumäärä e_i, joka kuvaa täysin kotitalouden tarpeet.[24] Ekvivalenttiskaalan avulla voidaan määritellä ns. ekvivalenttitulo y_{ekv}, joka kuvaa yksikäsitteisesti kotitalouden hyvinvoinnin tason. Muodollisesti se on määritelty yhtälönä

$$y_{ekv} = y_i/e_i, \tag{91}$$

jossa y_i on kotitalouden saamat tulot. Mitä suurempi kotitalouden ekvivalenttitulo on, sitä suurempi on kotitalouden hyvinvointi, tarpeista riippumatta.

Tarpeisiin vaikuttavat monet tekijät, mutta tyypillisesti huomioidaan ainakin kotitalouden koko ja jäsenten ikärakenne. Voidaankin hyvin ajatella, että lapset eivät vaadi samaa tulotasoa kuin aikuiset saavuttaakseen saman hyvinvoinnin tason. Samoin on laita kenties eläkeikäisilläkin, kunnes lääkekulut kasvavat suuriksi. Lisäksi, kuten edellä mainittiin, kotitaloudessa on mittakaavaetuja. Yksi televisio ja yksi jääkaappi riittävät useille kotitalouden jäsenille ja siten jokainen ei tarvitse omaansa. Näin saman hyvinvoinnin taso saavutetaan pienemmällä tulolla, kuin jos jokainen kotitalouden jäsen eläisi yksin omassa taloudessaan. Siten yli yhden henkilön kotitalouksille päteekin $e_i < h_i$, eli ekvivalenssiskaala on pienempi kuin kotitalouden jäsenten määrä. Lisäksi tyypillisesti tehdään implisiittinen oletus, jonka mukaan tulot jakautuvat kotitalouden jäsenille tasaisesti.

Ekvivalenssiskaalan määrittäminen on laaja ongelma ja kirjallisuudessa on tarjolla lukuisia erilaisia skaaloja joista valita.[25] Tässä tutkimuksessa on käytetty ns. OECD-skaalaa, jossa ensimmäinen aikuinen saa arvon yksi, sitä seuraavat aikuiset (ikä yli 17 vuotta) lisäävät skaalaa luvulla 0,7 ja lapset (alle 18 vuotta) luvulla 0,5. OECD-skaalaa on muutettu tämän määritelmän jälkeen, mutta tämä skaala on valittu, jotta tulokset olisivat vertailukelpoisia Suomessa aiemmin tehtyjen tutkimusten kanssa. Oletamme siis, että käytetty skaala on riittävän hyvä approksimaatio kotitalouden todellisista mittakaavaeduista.[26]

Ekvivalenssiskaalan valinnan vaikutuksia testataan käyttämällä kahta vaihtoehtoista skaalaa: modifioitua OECD-skaalaa sekä Atkinsonin neliöjuuriskaalaa. Modifioitu

[24] Huomaa, että ekvivalenttiskaalan arvo ei tyypillisesti ole kokonaisluku.

[25] Hagfors (1988) on esittänyt hyvän yleiskatsauksen ekvivalenssiskaaloihin ja skaalan estimointiin Suomen aineistolla.

[26] Suoniemi ja Sullström (1995, luku 7) ovat tarkastelleet OECD-skaalaa Suomen aineistolla.

OECD-skaala on samantapainen kuin edellä kuvattu, mutta tässä versiossa ensimmäinen aikuinen saa suhdeluvun 1, muut aikuiset (ikä yli 13 vuotta) suhdeluvun 0,5 ja lapset (ikä alle 14) suhdeluvun 0,3. Atkinsonin neliöjuuriskaalassa puolestaan ekvivalenssiskaala on kotitalouden henkilömäärän neliöjuuri.

Oikean ekvivalenssiskaalan löytäminen ja käyttäminen ei vielä ratkaise jakaumavertailun ongelmia. Skaalan lisäksi on löydettävä myös oikea painotus erikokoisille kotitalouksille, jotta tarkastelu todella tehtäisiin jaksossa 2.1 esiteltyjen periaatteiden mukaisesti. Tutkimuksissa on käytetty monenlaisia painotuksia riippuen tutkimusasetelmasta ja mielenkiinnon kohteesta. Taulukkoon 2 on koottu yleisimmät tulo- ja talousyksikkökombinaatiot. Tulotyypillä tarkoitetaan tässä yhteydessä tuloa, jonka suhteen kotitalouksia vertaillaan. Esimerkiksi jos tulo jaetaan kotitalouden jäsenten lukumäärällä, tarkastellaan silloin tulotyyppiä per capita tulo eli tuloa per jäsen. Tällöin tarkastelun kohteena oleva yksikkötyyppi on yksilö, joka edellä mainitun per capita tulon saa.

Taulukko 2 Menetelmiä kotitalousvertailuun. y on kotitalouden tulo, h kotitalouden jäsenten määrä ja e kotitalouden ekvivalenttiskaala eli -aikuisten lukumäärä.

Menetelmä	Tulo	Paino	Tuloyyppi	Yksikkötyyppi
1	y	1	kotitalouden tulot	kotitalous
2	y/h	h	per capita tulo	kotitalouden jäsen
3	y/e	h	ekvivalenttitulo	kotitalouden jäsen
4	y/e	e	ekvivalenttitulo	ekvivalenttiaikuinen

Menetelmässä 1 vertaillaan suoraan kotitalouksia. Näin toimittaessa ei huomioida lainkaan eroja kotitalouksien rakenteissa, vaan vertailukohtana ovat kotitalouden tulot ja yksikkönä kotitalous kokonaisuutena. Hyvinvoinnin kannalta oleellista tietoa jää käyttämättä, sillä menetelmän mukaan esimerkiksi kuuden hengen perhe saa saman elintason kuin yhden hengen talous, kun talouksien tulot ovat samat. Suurten talouksien hyvinvointia siis liioitellaan. Menetelmässä 2 kotitalouden jäsenelle annetaan kotitaloutensa per capita tulo. Oletuksena on siis, ettei kotitaloudessa ole lainkaan mittakaavaetuja. Menetelmä ei myöskään siis soveltune hyvinvointitarkasteluihin, sillä nyt em. kuuden hengen talous vaatii kuusinkertaiset tulot yksin asuvaan verrattuna, jotta hyvinvointi olisi talouksilla sama. Suurien talouksien hyvinvointia siis aliarvioidaan, jos kotitaloudessa on mittakaavaetuja. Menetelmä 3 huomioi mittakaavaedut ekvivalenssiskaaloin ja painotus tehdään yksilötasolla. Tämä menetelmä on yleisesti käytetty, eikä olekaan enää ilmeistä onko menetelmässä hyvinvointiharhaa erikokoisten talouksien välillä. Ebert (1997) on kuitenkin osoittanut, että jos hyvinvointivertailu perustellaan hyvinvointifunktioita ja ekvivalenssiskaaloja käyttäen, on menetelmä 4 ainoa oikea tapa vertailla erilaisten kotitalouksien muodostamia jakaumia. Tällöin siis vertaillaan kotitalouksien (laskennallisten) ekvivalenttiaikuisten jakaumaa.

Ekvivalenttiaikuisten jakauman käyttöä on helppo vastustaa sillä argumentilla, että jokaisen yksilön hyvinvointi on yhtä arvokas, eikä siksi voida antaa kotitaloudelle sen henkilölukua pienempää painoarvoa. Perustelut ovat kuitenkin kestämättömiä, sillä jokainen yksilö otetaan lähtötilanteessa (yhtälö (92) jaksossa 5.3.2) huomioon. Ekvivalenttiaikuisten jakaumaan päädytään, koska se on matemaattinen seuraus hyvinvoinnin lisääntymisestä progressiivisten tulonsiirtojen myötä. Ekvivalenttiaikuisten jakauma on puhtaasti teoreettinen apuväline, joka mahdollistaa jakaumien väliset hyvinvointivertailut Atkinsonin (1970) esittämin perustein, vaikka tarkastelujen lähtökohtana on heterogeenisten kotitalouksien muodostama aineisto.

Ekvivalenssiskaaloilla on siis jakaumatarkastelussa kaksi roolia: Ne ottavat huomioon erot kotitalouksien tarpeissa ja mahdollistavat siten hyvinvointivertailut erilaisten kotitalouksien välillä. Toisaalta hyvinvointitarkastelu ekvivalenssiskaaloja käyttäen johtaa tarkastelun teoreettisten ekvivalenttiaikuisten jakaumaan, jossa kotitalous saa painokseen ekvivalenssiskaalan arvon. Koska ekvivalenssiskaala saa näin kaksinkertaisen merkityksen, on sen kyky mallintaa kotitalouden tarve-eroja entistä suuremmassa roolissa. Tästä syystä empiirisessä analyysissä on syytä tehdä herkkyysanalyysi kokeilemalla erilaisia ekvivalenssiskaaloja. Jos tulokset vaihtelevat merkittävästi skaalojen vaihdon myötä, ovat tulokset epäluotettavia ja tällöin on syytä pyrkiä estimoimaan ekvivalenssiskaalat uudestaan aineiston olosuhteita vastaaviksi.

5.3.2 Ekvivalenssiskaalojen käytön teoreettinen johtaminen

Koska edellä mainittu Ebertin (1997) tulos on tärkeä, esitetään tässä tuloksesta pääkohdat. Lähtökohtana on siis jakaumaan kuuluvien yksilöiden kokeman hyvinvoinnin keskiarvo

$$W(\mathbf{y}) = a \sum_{i=1}^{k} \sum_{j=1}^{m_i} h_i U_i(y_{ij}), \tag{92}$$

jossa a on keskiarvoistava vakio. Erilaisia kotitalousryhmiä on k kpl, ryhmässä i on m_i kotitaloutta ja kaikissa ryhmään i kuuluvissa kotitalouksissa on h_i jäsentä. Lisäksi U_i on ryhmään i kuuluvan kotitalouden kunkin yksilön kokema hyöty, kun kotitalouden tulot ovat y_{ij}. Toisin kuin tässä tutkimuksessa Ebert käyttää artikkelissaan kotitalouksien hyötyfunktioita. Tulokset ovat molemmilla lähestymistavoilla samat. Yhtälö (92) on yhtälön (2) ryhmitelty vastine, joten myös nyt hyötyfunktio on tulkittava jonkin ulkopuolisen tarkkailijan määräämäksi tai sen sijaan on tehtävä voimakkaita oletuksia hyödyn kardinaalisuudesta ja hyötyfunktioiden samanlaisuudesta ryhmien sisällä. On tärkeää huomata, että nyt lähtötilanteessa sallitaan ryhmien välille erilaiset hyötyfunktiot. Tavoitteena on nyt ekvivalenssiskaalojen avulla päästä tilanteeseen, jossa kaikilla ryhmillä on sama hyötyfunktio, kuten yhtälössä (2) on.

Oletetaan nyt, että tunnemme ryhmien ekvivalenssiskaalat e_i. Skaala on ryhmät määrittelevien parametrien funktio

$$e_i = e(h_i, \mathbf{x_i}), \tag{93}$$

jossa $\mathbf{x_i}$ on rakennemuuttujien vektori, joka siis osaltaan määrittelee ryhmät. Esimerkiksi OECD-skaalassa ryhmän määrittelee kotitalouden jäsenmäärä ja ikärakenne. Tavoitteena on löytää *yksi* hyötyfunktio U, jolla voidaan kuvata kaikkien eri ryhmien yksilöiden kokema hyöty, käyttäen apuna ekvivalenssiskaalaa eli

$$W(\mathbf{y}) = a' \sum_{i=1}^{k} \sum_{j=1}^{m_i} \alpha_i U(y_{ij}/e_i), \tag{94}$$

jossa α_i on ryhmän i kotitalouksien saama paino, kun kaikkien ryhmien yksilöt saavat saman hyötyfunktion, argumenttinaan y_{ekv}. Kaikkien hyötyfunktioiden oletetaan olevan kahdesti jatkuvasti differentioituvia, konkaaveja ja aidosti kasvavia ($U' > 0$, $U'' \leq 0$).

Kuten aiemmin on esitetty, oletamme vähenevien tuloerojen (eriarvoisuuden) lisäävän hyvinvointia, jos kokonaistulot pysyvät ennallaan. Tämä johti eriarvoisuusmittojen yhteydessä siirtoperiaate-aksioomaan (ks. jakso 3.1). Nyt tehtävässä tarkastelussa on huomioitava kotitalouksien erilaiset tarpeet ja siten siirtoperiaatettakin on hieman yleistettävä. Nyt nimittäin on huomioitava ekvivalenttitulon vaikutus hyvinvointiin, mutta samalla myös se, että tulonsiirto tapahtuu kuitenkin aidoissa tulossa. Kummassakin esityksessä kuvataan tulonsiirtoa, joka vähentää tuloeroja. Ebert (1997) esittää periaatteen ekvivalenttitulolle muodossa

Ominaisuus PS(e), Progressiivinen siirto. Pienen tulomäärän $\epsilon > 0$ siirto ryhmän i kotitaloudesta k, ryhmän j kotitalouteen l, s.e. $y_{ik} \to y_{ik} - \epsilon$ ja $y_{jl} \to y_{jl} + \epsilon$, parantaa hyvinvointia, jos $y_{jl}/e_j < y_{ik}/e_i$ ja $(y_{jl} + \epsilon)/e_j < (y_{ik} - \epsilon)/e_i$.

Oletetaan nyt yhtälön (92) toteuttavan ominaisuuden PS(e). Tehdään ominaisuuden kuvaama tulonsiirto, jolloin pätee

$$h_i U_i(y_{ik} - \epsilon) + h_j U_j(y_{jl} + \epsilon) > h_i U_i(y_{ik}) + h_j U_j(y_{jl}),$$

ryhmittelemällä ja jakamalla puolittain ϵ:lla saadaan erotusosamäärä, jonka raja-arvo $\epsilon \to 0$ on

$$h_i U_i'(y_{ik}) < h_j U_j'(y_{jl}),$$

kun $y_{jl}/e_j < y_{ik}/e_i$. Jos siirto tehdäänkin toisinpäin ja siirto toteuttaa edelleen ehdon $PS(e)$ eli nyt toisen kotitalouden ekvivalenttitulot ovat suuremmat, saadaan

$$h_i U_i'(y_{ik}) > h_j U_j'(y_{jl}),$$

kun $y_{jl}/e_j > y_{ik}/e_i$. Koska U on kahdesti differentioituva, voidaan todeta, että

$$h_i U_i'(y_{ik}) = h_j U_j'(y_{jl}), \tag{95}$$

kun $y_{jl}/e_j = y_{ik}/e_i$. Integroimalla saatu yhtälö saadaan tulos[27]

$$U_i(y_{ik}) = \frac{h_j\, e_i}{h_i\, e_j} U_j \left(\frac{e_j}{e_i} y_{ik} \right) + U_0. \tag{96}$$

Kun ryhmäksi j valitaan yhden hengen taloudet, saadaan muut ryhmät normitettua

$$U_i(y_{ik}) = \frac{e_i}{h_i} U(y_{ik}/e_i), \tag{97}$$

sillä tällä valinnalla $h_j = e_j = 1$. Lisäksi integrointivakio U_0 on valittu nollaksi[28] ja merkintä $U_j \equiv U$ on otettu käyttöön. Sijoittamalla yhtälö (97) yhtälöön (92) saadaan yhtälö

$$W(\mathbf{y}) = a \sum_{i=1}^{k} \sum_{j=1}^{m_i} e_i U(y_{ij}/e_i), \tag{98}$$

[27] $h_i \int U_i'(y)dy = h_j \int U_j'(ay)dy = h_j \frac{1}{a} \int U_j'(x)dx$, jossa $a = \frac{e_j}{e_i}$

[28] Vakion asettaminen nollaksi on perusteltua, koska haluamme vertailla hyvinvointifunktioiden suuruusjärjestystä, jolloin vakioilla ei ole merkitystä.

joka on haluttua muotoa (94). Tärkeää on huomata, että hyvinvointifunktiota painotetaan ekvivalenttien aikuisten määrällä e_i (α_i yhtälössä (94)). Jotta siis hyvinvointitarkastelu tehtäisiin hyötyfunktioihin perustuen, on siirryttävä tarkastelemaan (laskennallisten) ekvivalenttiaikuisten jakaumaa.

Näin muotoillun hyvinvointifunktion optimissa kaikkien ryhmien yksilöt saavat saman hyödyn. Jos paino α_i olisi kotitalouden jäsenten lukumäärä h_i (edellä mainitussa menetelmässä 3), niin optimitilanteissa suuren kotitalouden jäsenet saisivat suuremman hyödyn kuin pienemmän kotitalouden jäsenet (Ebert 1997), mikä ei tunnu hyväksyttävältä. Ebert (1999) esittelee ekvivalenttiskaaloihin perustuvan hyvinvointiteorian perusteellisesti.

On syytä huomata, että lopputulos on intuitiivisesti ilmeinen: jos kotitalous halutaan kuvata ekvivalenttiaikuisten joukkona, niin tällöin tämän joukon koko on kotitalouden luonnollinen painokerroin. Koska kotitalous on muutettu standardi-mittayksiköihin, on siihen liittyvien painokertoimien vastattava tätä samaista mittayksikköä.

6 Tuloerot ja eriarvoisuus Suomessa vuosina 1985–2001

Tässä luvussa tutkitaan tuloeroja Suomessa sekä tulonsiirtojen vaikutusta tuloeroihin edellä määritettyjen periaatteiden mukaisesti, edellisessä luvussa kuvatuilla kulutustutkimusten aineistoilla. Tulonsiirtojen hyvinvointivaikutukset olisivat luonnollisesti myös mielenkiintoisia, mutta niiden tutkiminen tämän kaltaisella aineistolla on mahdotonta. Nettotulonsiirrot nimittäin laskevat kotitalouksien tulojen keskiarvoa, jolloin yhteiskunnan hyvinvointi laskee, jollei eriarvoisuus laske samalla tarpeeksi voimakkaasti. Menetetty tulo ei kuitenkaan poistu yhteiskunnasta, vaan niillä tuotetaan julkisen sektorin tarjoamia palveluja, joiden käyttöön ja siten tuottamaan hyvinvointiin ei kulutustutkimus anna tietoa.

Edellä esitettyjä menetelmiä käyttäen tutkitaan, kuinka tuloerot ja eriarvoisuus ovat Suomessa viime vuosina (1985–2001) kehittyneet. Lisäksi tarkastellaan koko verorasituksen jakautumista kotitalouksien kesken. Ensimmäiseksi on kuitenkin selvitettävä, miten eri eriarvoisuusmittarit soveltuvat käytössä olevan aineiston tarkastelemiseen. Tämän vuoksi aluksi suoritetaan tarkasteluja, jotka tuovat esiin mahdolliset ongelmat. Kun sopiva mitta on valittu, käytetään sitä eri tekijöiden vaikutuksen selvittämiseen.

Kuten aiemmin on mainittu, eriarvoisuusmittojen ongelmana on parametrin valinta. Koska valinta on lähes mielivaltainen, käytetään mittareita tässä tutkimuksessa siten, että kaikki mielekkäät parametriarvot tulevat tarkastelluksi. Tämä toteutetaan tutkimalla mittojen arvoja graafisesti mitan parametrin funktiona.

6.1 Tuloeriarvoisuus

Tarkastellaan ensin käytettävissä olevan tulon (nettotulon) jakaumaa. Oletuksena on, että nettotulo kuvaa parhaiten taloudellista hyvinvointia, joten tarkasteluille voidaan antaa eriarvoisuustulkinta.[29] Tarkastelut tehdään kaikilla kolmella luvussa 3 esitellyllä mitalla.

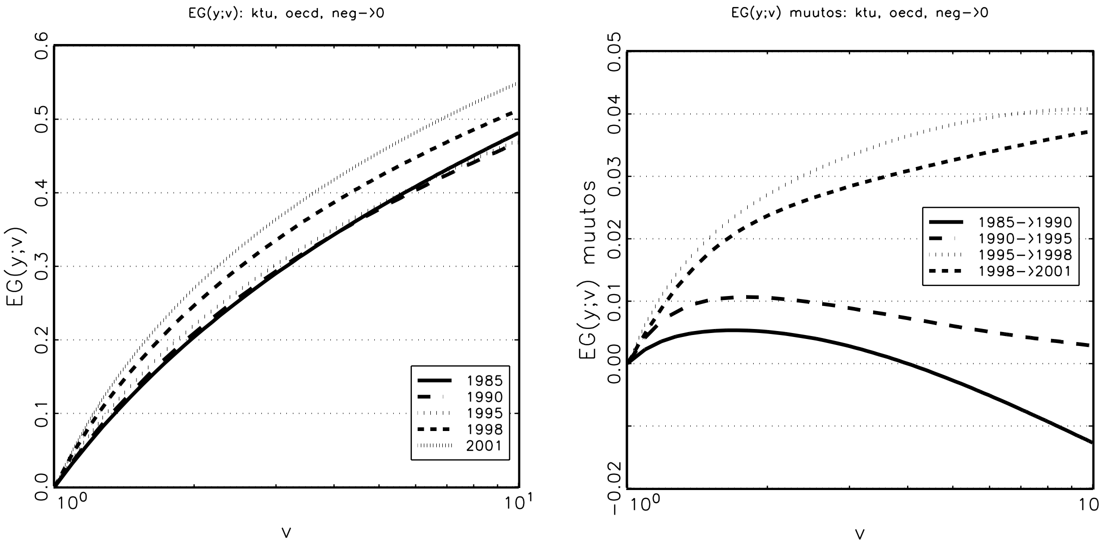
6.1.1 Laajennettu Gini-kerroin

Kuvassa 11 on esitetty laajennetun Gini-kertoimen arvot sekä aikamuutokset parametrinsa funktiona.[30] Havaitaan, että mittarin arvo kasvaa parametrin kasvaessa. Tämä johtuu mittarin ominaisuuksista: samaa tilannetta kuvattaessa eri parametrin arvoilla, suurempi parametrin arvo tuottaa suuremman mitan arvon.[31] Lisäksi havaitaan, että

[29] Otsikon termi tuloeriarvoisuus kuvaa eroa myöhemmin tarkasteltavaan kulutusjakaumaan pohjautuvaa eriarvoisuutta.

[30] Kuvan yllä oleva otsake kertoo, mitkä valinnat aineistoa koskien on tehty. Kaikki havainnot ovat käytössä, mutta negatiiviset havainnot on muutettu nolliksi. Ekvivalenssiskaalana on OECD-skaala, joka on oletusskaala tehdyissä tarkasteluissa.

[31] Vain erikoistilanteissa näin ei käy, esimerkiksi jos tulot ovat tasaisesti jakautuneet, on mitan arvo aina 0.



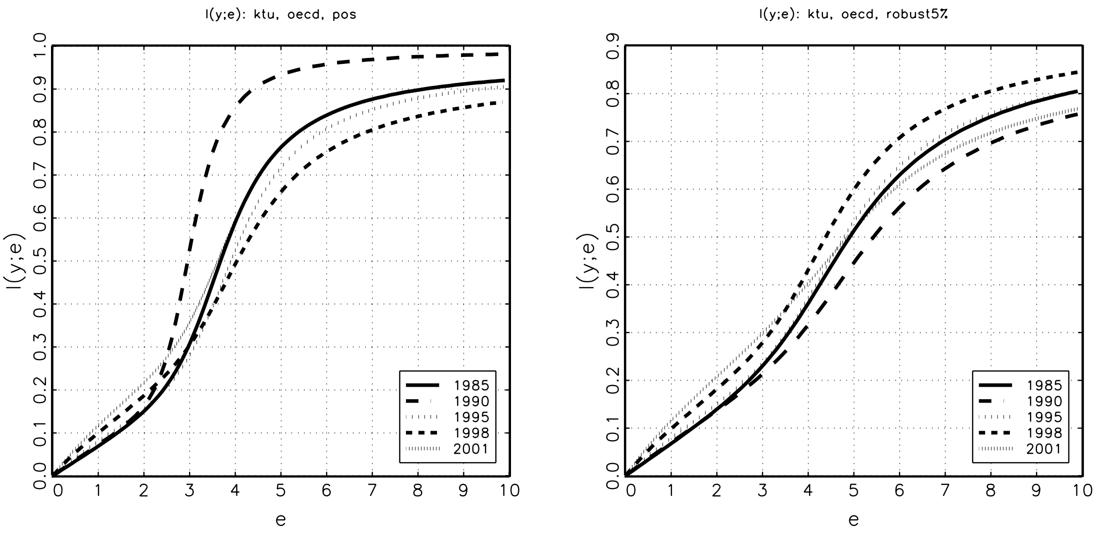
Kuvio 11 Laajennettu Gini-kerroin parametrinsa funktiona ja sen aikamuutokset, kun tulokäsitteenä on käytettävissä olevat tulot.

pienillä parametrin arvoilla tuloerot ovat kasvaneet kaikkien mittausten välillä. Suurilla parametrin arvoilla muutos vuodesta 1985 vuoteen 1990 on sen sijaan ollut tuloeroja kaventava ja siirryttäessä lamaan (1990 → 1995) tuloerot ovat kasvaneet suurten parametrien arvoilla vähemmän kuin pienillä arvoilla.[32] Tämän jälkeen muutos on ollut suurinta suurilla parametrin arvoilla. Koska pienet parametrin arvot korostavat suhteellisesti suurten tulojen kaihtamista ja suuret vastaavasti pienten tulojen kaihtamista, voidaan kuvasta päätellä, että suhteellisesti suurituloisten osuus jakaumassa on kasvanut kaikkien mittausten välillä ja että vuoteen 1995 asti suhteellisesti pienituloisten määrä pysyi samalla tasolla, jonka jälkeen myös niiden osuus on kasvanut.

6.1.2 Muut mitat

Vastaava tarkastelu voidaan tehdä Atkinsonin indeksillä sekä yleistetyn entropian mitalla. Jotta parametrialueet olisivat vapaasti käytettävissä, tutkitaan vain positiivisten käytettävissä olevien tulojen jakaumaa. Koska mitat ovat osassa parametrialueestaan hyvin herkkiä pienille tai suurille havainnoille, ajaudutaan helposti ongelmiin äärihavaintojen kanssa. Kuvissa 12 ja 13 on esitetty edellä mainittujen mittojen käyrät kahdessa eri tapauksessa. Vasemman puoleisissa kuvissa mittojen laskuun on käytetty kaikkia positiivisia havaintoja. Oikealla jakaumaa on käsitelty, jotta äärihavainnot eivät turmelisi mittojen arvoja. Atkinsonin indeksin tapauksessa riittää kontrolloida pienimpiä havaintoja, sillä suurilla parametrin arvoilla mitta lähestyy rawlsilaista hyvinvointi-ideaa, jolloin pienin arvo määrää kokonaan mitan arvon. Tämä ei tietenkään sovellu otosaineiston tapaukseen. Rajaus on tässä tapauksessa tehty pudottamalla kaikki havainnot, joiden arvot ovat alle 5 % keskiarvotulosta. Yleistetyn entropian mitta on herkkä myös suurille havainnoille. Suurilla havainnoilla käsittely tehtiin poistamalla 10

[32] Käytettäessä modifioitua OECD-skaalaa tai Atkinsonin neliöjuuri-skaalaa, laskevat muutokset suurten parametrien arvoilla myös tässä tapauksessa negatiivisiksi.



Kuvio 12 Atkinsonin indeksi laskettuna käytettävissä olevien tulojen jakaumalle. Vasemmassa kuvassa on käytetty kaikkia positiivisia havaintoja ja oikean puoleisessa jakauma on katkaistu alhaalta tekstissä kuvatulla tavalla.

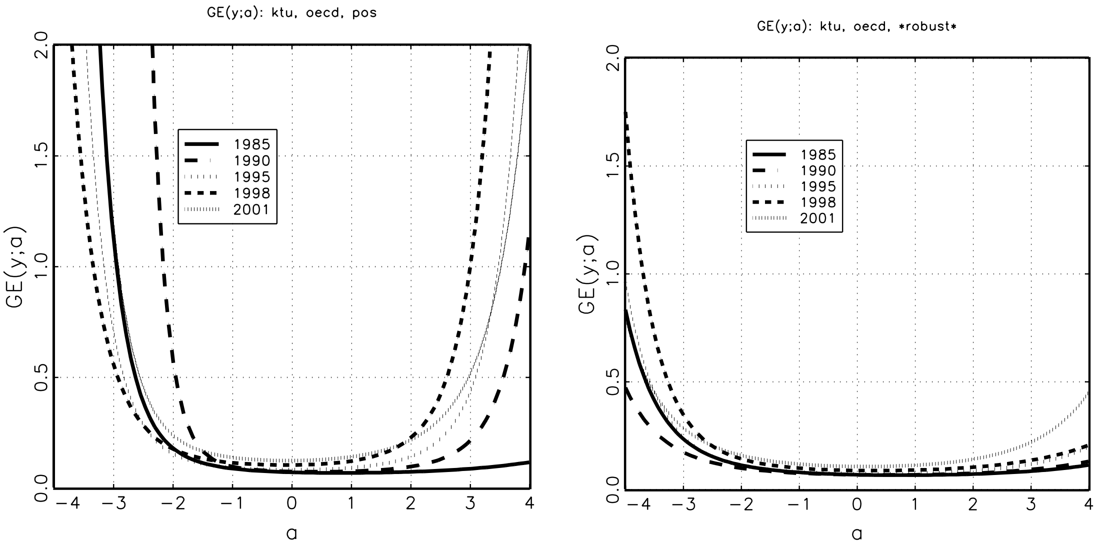
suurinta havaintoa tarkasteluista. Vaikutukset ovat huomattavat, kuten kuvista nähdään. Kuvien perusteella voidaan arvioida, että Atkinsonin indeksin arvot parametrin arvoilla $\epsilon > 2$ ovat epäluotettavia ja pyrkivät kuvaamaan pienimmän havainnon suuruutta. Tämä vastaa yleistetyn entropian parametrialuetta $\alpha < -1$. Yleistetyn entropian arvot ovat epäluotettavia myös suurilla parametrin arvoilla ($\alpha > 2$), sillä muutos on dramaattinen, kun 10 suurinta havaintoa poistettiin tutkitusta aineistosta.

Kuvassa 14 on edellä kuvatuista ongelmista huolimatta esitetty tuloerojen muutokset.[33] Kuvassa on käytetty katkaistuja jakaumia, mutta tästä huolimatta itseisarvoltaan suuret parametrin arvot ovat epäluotettavia. Kuvista havaitaan myös se, että yleistetty entropia järjestää jakaumat identtisesti Atkinsonin indeksin kanssa, kun entropiaparametri $\alpha = 1 - \epsilon$, jossa ϵ on Atkinsonin indeksin parametri.[34]

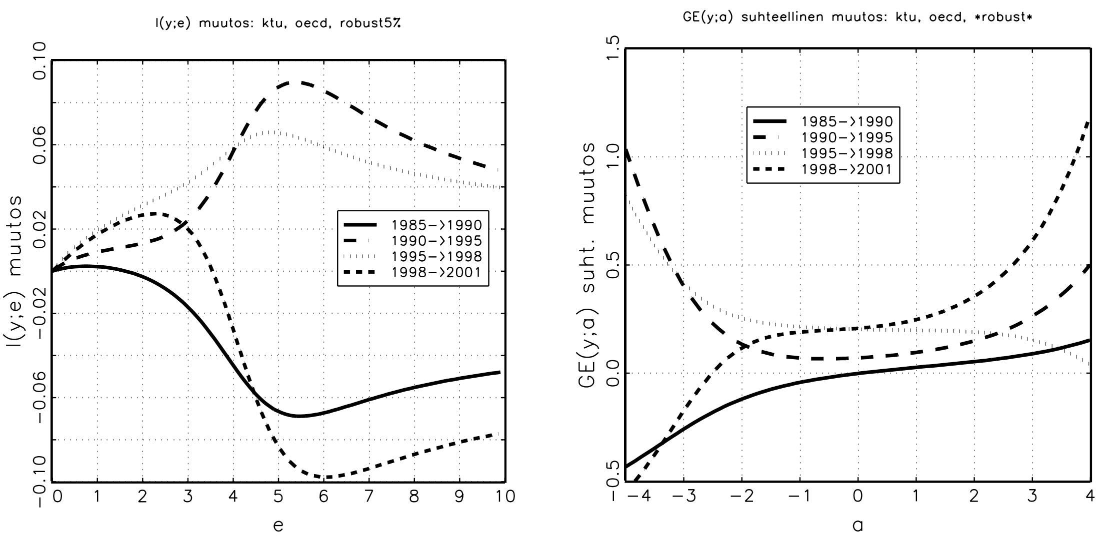
Keskeiset tulokset ovat näillä mitoilla identtiset laajennetulla Gini-kertoimella saatuihin tuloksiin nähden, vaikka jakaumat poikkeavat äärihavaintojen käsittelyn osalta. Atkinsonin indeksin luotettavalla alueella ($\epsilon < 2$) havaitaan samanlainen käytös kuin edellä: jakaumien järjestys on sama ja matalien tulojen painotus järjestää vuoden 1990 jakauman vuoden 1985 jakaumaa vähemmän eriarvoiseksi. Yleistetyn entropian arvioitu luotettava väli on $[-1, 2]$. Välillä $[-1, 1]$ järjestys on määritelmällisesti identtinen Atkinsonin indeksin kanssa, joten edellä mainittu pätee myös nyt. Kun parametri $\alpha > 1$, niin mitta ei enää pidä nollahavaintoja pieninä ja siten painotus korostaa enemmän suurten tulojen osuutta tuloeroissa. Tällöin mitan mukaan tuloerot ovat kaikkien osa-aineistojen välillä kasvanut eli mitta järjestää aineistot kuten laajennettu Gini-kerroin. Järjestyksen samanlaisuudesta, poikkeavien havaintojen aiheuttamista sekä nolla-havaintoihin liittyvistä ongelmista johtuen, jatkotarkasteluissa käytetään vain laajennettua Gini-kerrointa.

[33] Yleistetyn entropian tapauksessa on paras kuvata suhteellisia muutoksia.

[34] Pieniä eroja voi syntyä siitä, että aineistoja on käsitelty suurten havaintojen osalta eri tavalla.



Kuvio 13 Yleistetyn entropian mitta laskettuna käytettävissä olevien tulojen jakaumalle. Vasemmassa kuvassa on käytetty kaikkia positiivisia havaintoja ja oikean puoleisessa jakauma on katkaistu alhaalta sekä ylhäältä tekstissä kuvatulla tavalla.



Kuvio 14 Atkinsonin indeksin sekä yleistetyn entropian muutokset parametrien funktiona. Tutkittu jakauma on tekstissä mainitulla tavalla katkaistu käytettävissä olevien tulojen jakauma.

6.1.3 Ekvivalenssiskaalojen vaikutus

Kuten edellisessä luvussa mainittiin, on rakenteeltaan poikkeavat kotitaloudet tehtävä yhteismitalliseksi, jotta tuloerotarkastelut ovat mielekkäitä. Ongelma ratkaistiin tässä työssä ekvivalenssiskaaloja käyttäen, mutta skaalan valinta on lähes mielivaltainen päätös. Tämän vuoksi on syytä tarkastella, kuinka paljon tulokset ovat riippuvaisia skaalan valinnasta.

Kuvissa 15 ja 16 on esitetty käytettävissä olevien tulojen laajennetun Gini-kertoimen arvot, kun ekvivalenssiskaalana on käytetty modifioitua OECD ja Atkinsonin neliöjuuri-skaalaa (vertaa kuvaan 11). Havaitaan, että saadut tulokset eivät juuri muutu. Jakaumien järjestys pysyy samana lähes kaikilla parametrien arvoilla: vain suurilla parametrin arvoilla vuoden 1995 jakauman tuloerot tulevat vuoden 1990 jakaumaa pienemmiksi. Erot ovat kuitenkin melko pieniä eivätkä vaikuta merkittävästi muutosten kokonaiskuvaan, eli vuosien 1985 ja 1995 välinen tuloerojen kasvu on lähinnä suurten tulojen osuuden kasvua, kun taas siitä eteenpäin tuloerojen kasvu johtuu niin suurten kuin pientenkin tulojen määrän kasvusta.

Voidaankin sanoa, että kunhan käytetään mielekkäästi muodostettua ekvivalenssiskaalaa, on skaalan valinta johtopäätösten suhteen *harmiton* yksinkertaistus (Vartia 1988). Skaalan käyttö tarve-erojen huomioimisessa on siis siinä mielessä hyvä approksimaatio, ettei sen valinta mielekkäistä vaihtoehdoista juuri vaikuta tuloksiin. Jatkossa tarkasteluihin käytetään vain OECD-skaalaa.

6.2 Eriarvoisuuden muutosten lähteet

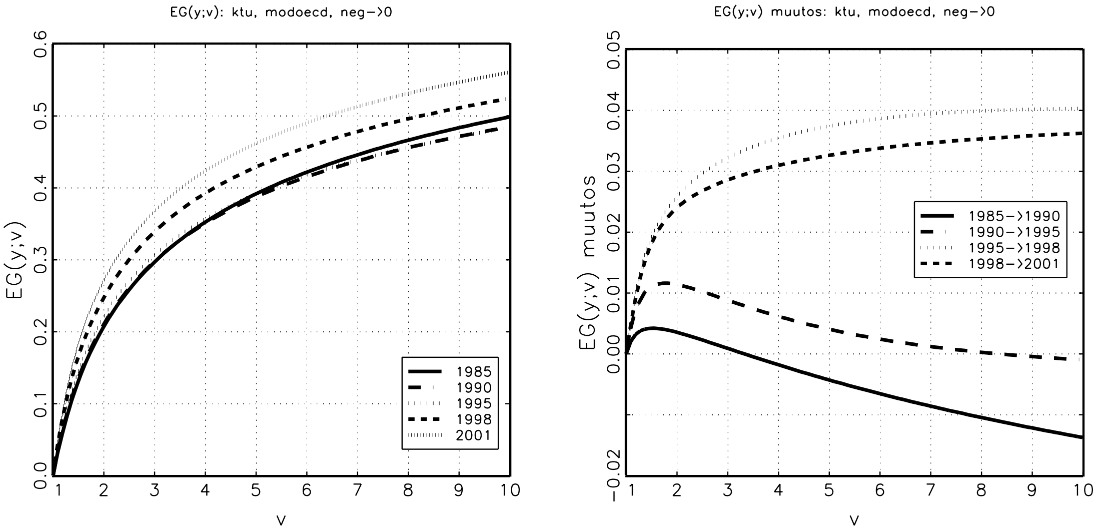
Edellisen jakson tulosten mukaan tuloerot ovat Suomessa kasvaneet edellisen vuosituhannen viimeisinä vuosina. Jotta paremmin ymmärtäisimme muutoksia ja niiden merkitystä, on syytä tarkastella tarkemmin, miten muutokset rakentuvat.

Käytettävissä olevien tulojen jakauman taustalla on tuotannontekijätulojen jakauma. Kun tuotannontekijätuloihin lisätään kotitalouden julkiselta sektorilta saamat ja sille maksamat tulonsiirrot, päädytään käytettävissä olevien tulojen jakaumaan (ks. taulukko 1 sivulla 40). Näiden tulokomponenttien muutoksia voidaan tarkastella yksitellen, jolloin muutoksen taustalla olevat tekijät hahmottuvat tarkemmin.

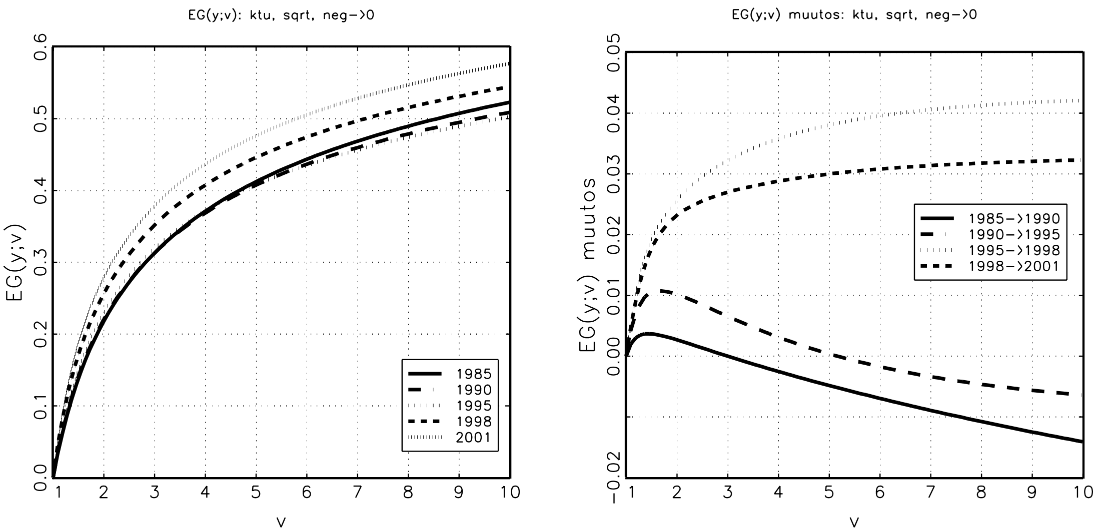
6.2.1 Tuotannontekijätulot

Tuotannontekijätulot muodostavat tässä tutkimuksessa tarkastelujen lähtökohdan. Tämä sen vuoksi, että tuotannontekijätulot koostuvat osin tulokomponenteista, jotka useissa kotitalouksissa saavat negatiivisia arvoja (yrittäjä- sekä omaisuustulot). Koska tässä työssä käsitelty teoria on tarkoitettu tilanteisiin, joissa tulot ovat ei-negatiivisia, joudutaan tuotannontekijätulojen rakentuminen muista tulokomponenteista sivuuttamaan ja jakauma otetaan siten annettuna.

Kuvassa 17 on esitetty tuotannontekijätulojen jakaumalle lasketut laajennetut Gini-

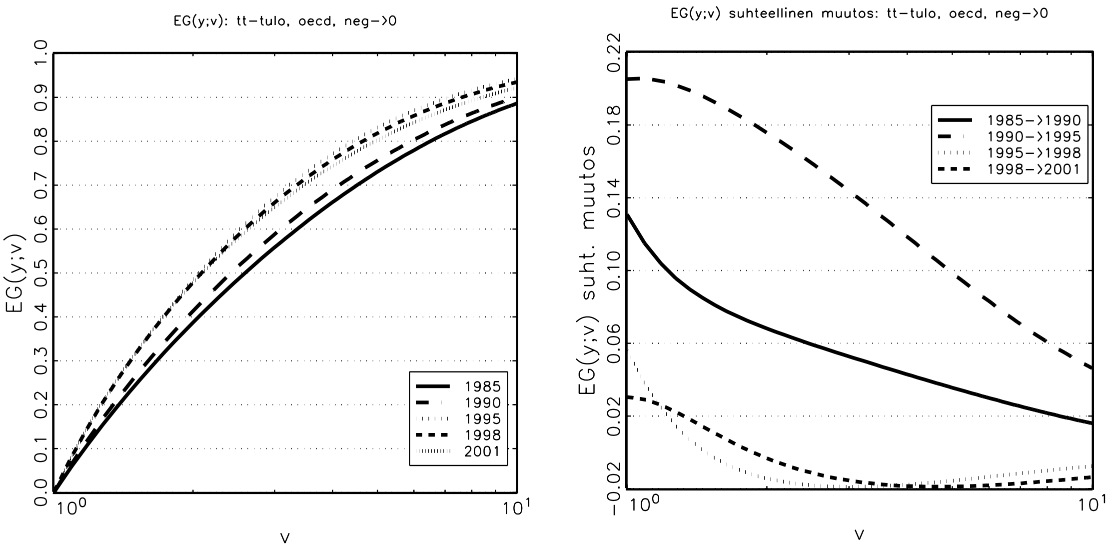


Kuvio 15 Laajennettu Gini-kerroin ja käytettävissä olevat tulot käytettäessä modifioitua OECD-ekvivalenssiskaalaa. Vertaa kuvaan 11. Huomaa, että kuvassa 11 vaaka-akselin asteikko on logaritminen.



Kuvio 16 Laajennettu Gini-kerroin ja käytettävissä olevat tulot käytettäessä Atkinsonin neliöjuuri-skaalaa. Vertaa kuvaan 11. Huomaa, että kuvassa 11 vaaka-akselin asteikko on logaritminen.

kertoimet sekä tämän suhteelliset aikamuutokset.[35] Havaitaan, että tuotannontekijätuloerojen kasvu on alkanut jo 80-luvun lopulla, mutta laman ja suurtyöttömyyden myötä tuloerot kasvoivat voimakkaasti 90-luvun alussa (1990 → 1995). Tämän jälkeen tuloerot ovat kaventuneet hivenen painotettaessa pieniä tuloja. Tämä lienee seurausta työllisyyden paranemisesta eli nollatulojen vähenemisestä. Toisaalta painotettaessa suuria tuloja tuloerot ovat jatkaneet kasvua, mutta aiempaa huomattavasti maltillisemmin.



Kuvio 17 Laajennettu Gini-kerroin parametrinsa funktiona ja sen aikamuutokset, kun tulokäsitteenä on tuotannontekijätulot.

6.2.2 Saadut tulonsiirrot

Tuotannontekijätulot ovat markkinoiden määrittämät ja muodostavat tulojakauman lähtökohdan. Jollei julkinen sektori vaikuttaisi toimillaan tulonjakoon, olisivat tuloerot edellä kuvatun kaltaiset. Näiden tulojen lisäksi kotitaloudet saavat valtiolta tulonsiirtoja sekä maksavat tälle tulonsiirtoja (käytännössä lähinnä välittömiä veroja). Jos saadut ja maksetut tulonsiirrot olisivat suoraan verrannolliset kotitalouden tuotannontekijätuloihin, eivät nämä erät muuttaisi suhteellisia tuloeroja. Näin ei kuitenkaan ole, ja siksi tuloeroissa havaitaan muutoksia.

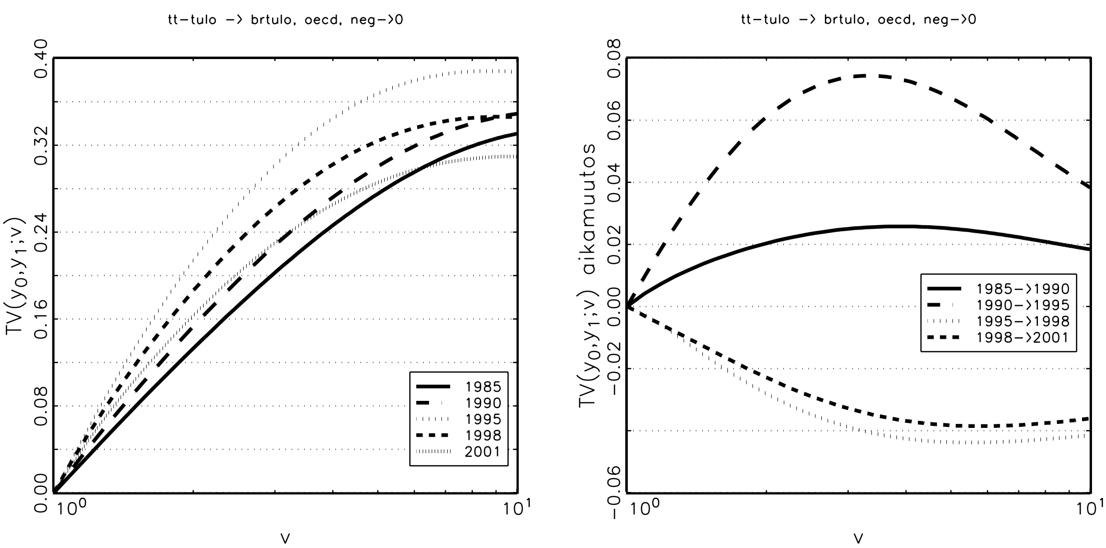
Julkisen sektorin toimien, saatujen ja maksettujen tulonsiirtojen, vaikutusta tuloeroihin voidaan tarkastella sen aiheuttamina muutoksina laajennetun Gini-kertoimen arvoon. Määritellään siis tarkasteluja varten tuloerojen vähenemisen mitta TV laajennetun Gini-kertoimen avulla

$$TV(y_0, y_1; \nu) = EG(y_0; \nu) - EG(y_1; \nu). \tag{99}$$

[35] Suhteelliset muutokset tuovat paremmin esiin muutokset pienillä parametrin arvoilla. Toisaalta mitan arvon kasvu parametrin kasvaessa johtaa siihen, että suhteellinen muutos suurilla parametrin arvoilla on maltillista.

Funktio TV on siis lähtöjakauman ja loppujakauman laajennettujen Gini-kertoimien erotus. Jakaumat ovat eri tulokomponenttien jakaumia *samalla* ajan hetkellä. Positiivinen TV:n arvo merkitsee siten suhteellisten tuloerojen pienenemistä ja negatiivinen kasvamista, kun siirrytään tulokäsitteestä toiseen. Vertailukohtana on siis tulonsiirto, joka on suoraan verrannollinen lähtöjakauman tuloihin. Tarkastelut tehdään kuten edellä esittämällä funktio TV graafisesti parametrinsa ν funktiona.

Kuvaan 18 on esitetty saatujen tulonsiirtojen TV-funktion arvot parametrin ν funktiona eli tuotannontekijätulojen ja bruttotulojen laajennettujen Gini-kertoimien erotus eri vuosina sekä näiden aikamuutokset. Kuvasta havaitaan, että saadut tulonsiirrot vähentävät tuloeroja. Näin mitattu tuloeroja vähentävä vaikutus voimistui siirryttäessä vuodesta 1985 laman huippuun 1995. Näin tulonsiirrot pystyivät vähentämään tuotannontekijätulojen kasvaneiden tuloerojen vaikutusta. Vuosituhannen lopulla saatujen tulonsiirtojen tuloja tasaava vaikutus on kuitenkin hieman vähentynyt, mikä havaitaan negatiivisista aikamuutoksista (oikean puoleinen kuva). Vuonna 2001 tuloerojen tasaus olikin suunnilleen samalla tasolla kuin vuonna 1985, vaikka tuloerojen lähtötaso oli korkeammalla. Samalla tulonsiirtojen vaikutuksen rakenne on muuttunut: painotettaessa voimakkaasti pieniä tuloja tulon tasaus oli vuonna 2001 vuotta 1985 heikompaa, mutta muulloin tätä voimakkaampaa.



Kuvio 18 Saatujen tulonsiirtojen vaikutus mitattuna TV-funktion (yhtälö (99)) avulla. Funktion arvot ovat tuotannontekijätulojen ja bruttotulojen Gini-kertoimien erotus.

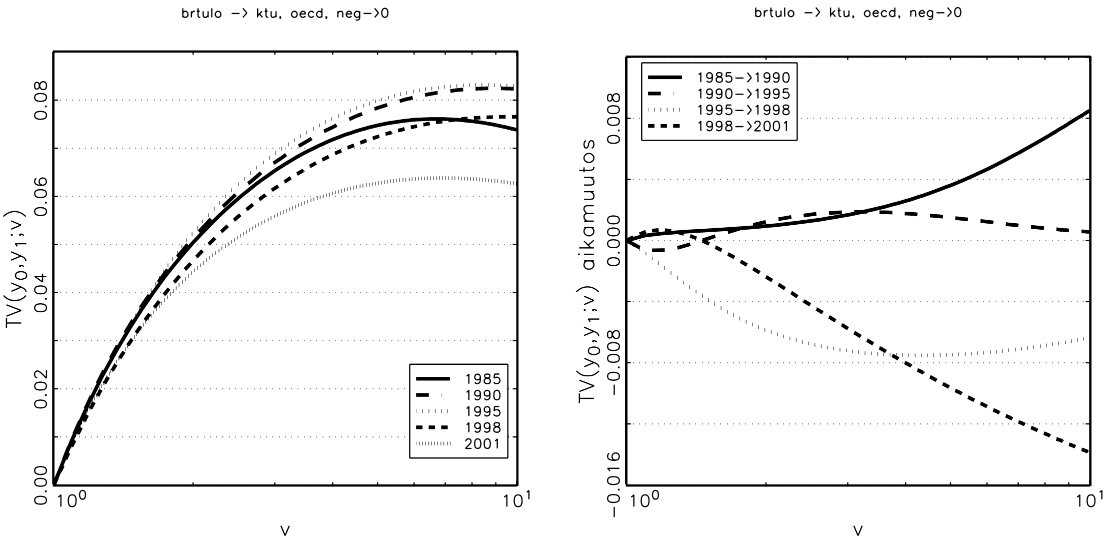
On syytä huomata, että käytetty menetelmä ei tuo esiin pelkän tulonsiirtojärjestelmän muutoksia, vaan lähtöjakauman *ja* tulonsiirtojärjestelmän yhteisen muutoksen. Lähtöjakaumassa tuloksiin vaikuttavat sekä tulojen rakenne että tuloerojen taso, joskin vuosina 1995–2001 lähtötaso on pysynyt liki muuttumattomana. Havaittujen muutosten taustalla olevien tekijöiden tarkempi selvittäminen vaatii lisätutkimusta.

6.2.3 Maksetut tulonsiirrot

Kun kotitaloudet ovat saaneet tulonsiirrot julkiselta sektorilta, joutuvat ne maksamaan tulonsiirtoja, lähinnä siis välittömiä veroja. Verojärjestelmä on Suomessa progressiivinen, joten myös välittömät verot pyrkivät vähentämään tuloeroja. Edellä käytettyä menetelmää voidaan soveltaa myös tässä tapauksessa.

Funktio TV kuvaa nyt verotuksen efektiivistä progressiota (esim. Musgrave ja Thin 1948) ja koostuu verotuksen tuloja tasaavasta vaikutuksesta sekä verotuksen aiheuttamasta tulojärjestyksen muutoksesta, ns. horisontaalisesta eriarvoisuudesta (esim. Lambert 2001, 238–242 ja Lerman ja Yitzhaki 1995). Suomessa tulojärjestyksen muutoksen vaikutus on maksettujen tulonsiirtojen osalta melko vähäistä (Riihelä ja Sullström 2004), joten havaitut TV:n arvot kuvastavat suurelta osin verojärjestelmän progressiivisuutta. Myös nyt on syytä muistaa, että aikamuutoksien taustalla on myös lähtöjakaumassa tapahtuneet muutokset eikä pelkästään verojärjestelmän muutokset.

Kuvassa 19 on esitetty maksettujen tulonsiirtojen TV-funktion arvot parametrin funktiona eli laajennetun Gini-kertoimen erotus siirryttäessä bruttotuloista käytettävissä oleviin tuloihin ja näiden ajalliset muutokset. Kuvasta havaitaan, että kotitalouksien



Kuvio 19 Maksettujen tulonsiirtojen vaikutus mitattuna TV-funktion (yhtälö (99)) avulla. Funktion arvot ovat bruttotulojen ja käytettävissä olevien tulojen Gini-kertoimien erotus.

maksamat tulonsiirrot vähentävät tuloeroja. Vaikutukset ovat noin viidesosa saatujen tulonsiirtojen vaikutukseen nähden.[36] Lisäksi havaitaan, että vaikutusten muutokset ovat olleet huomattavasti monipuolisempia kuin saatujen tulonsiirtojen tapauksessa (vrt. kuva 18). Vasemman puoleisesta kuvasta nähdään, että maksettujen tulonsiirtojen tuloeroja tasaava vaikutus on suurimmillaan vuosina 1990 ja 1995, jonka jälkeen vaikutus on vähentynyt. Siirryttäessä vuodesta 1985 vuoteen 1990 tuloerojen tasaan-

[36] Laajennetun Gini-kertoimen lähtötasoissa on suuri ero ja tämä vaikuttaa absoluuttisiin muutoksiin.

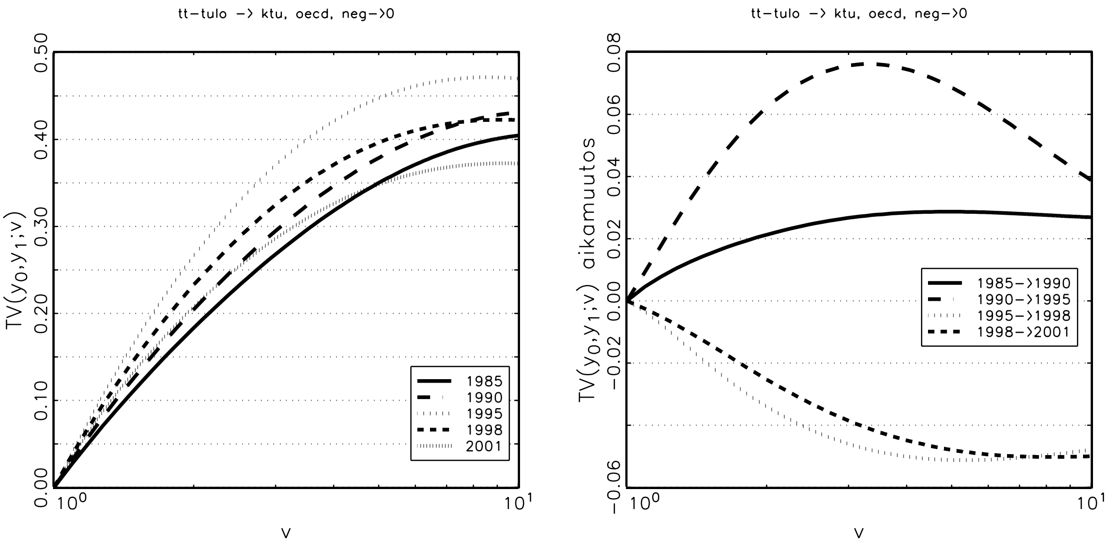
tumisen kasvua havaitaan kaikilla painotuksilla, mutta etenkin kun tarkastellaan pienituloisia. Siirryttäessä lamaan (1990 → 1995) painotettaessa suurituloisia havaitaan tuloerojen tasauksen vähenevän ja painotettaessa pienituloisia kasvavan, mutta muutokset ovat hyvin pieniä. Vuosituhannen loppupuolella maksettujen tulonsiirtojen tuloeroja tasaava vaikutus heikkeni etenkin siirryttäessä vuodesta 1995 vuoteen 1998, mutta tämän jälkeen muutos riippuu painotuksesta: painotettaessa suurituloisia, havaitaan tuloerojen tasoittumisen voimistumista, mutta painotettaessa pienituloisia, sen heikkenemistä.

6.2.4 Efektiiviset verot

Edellä on kuvattu julkisen sektorin toimien vaikutusta tulojakaumaan saatujen ja maksettujen tulonsiirtojen välityksellä. Jotta kokonaiskuva ei hämärtyisi, tarkastellaan vielä näiden vaikutusta yhtä aikaa eli julkisen sektorin kokonaisvaikutusta markkinaratkaisusta havaituksi eriarvoisuudeksi. Tällöin tarkastellaan ns. *efektiivistä verotusta*. Jos tulonsiirrot tulkitaan negatiivisena verona, saadaan kotitalouden efektiiviseksi veroksi

$$TE = T - S, \tag{100}$$

jossa T on kotitalouden maksamat tulonsiirrot ja S sen saamat tulonsiirrot. Jos tulonsiirtoja saadaan enemmän kuin maksetaan, ovat kotitalouden efektiiviset verot negatiiviset.



Kuvio 20 Efektiivisten verojen eli saatujen ja maksettujen tulonsiirtojen vaikutus tuloeroihin. TV-funktion arvot ovat tuotannontekijätulojen ja käytettävissä olevien tulojen Gini-kertoimien erotus.

Tulonsiirtojen kokonaisvaikutusta voidaan arvioida tarkastelemalla efektiivisten verojen vaikutusta. Tämä tapahtuu vertailemalla siirtymää tuotannontekijätuloista nettotuloihin. Kuvassa 20 on esitetty efektiivisen verotuksen vaikutus tulonjakoon. Kuvasta havaitaan, että efektiivinen verotus käyttäytyy pitkälti samoin kuin saatujen tulonsiirtojen tapaus. Tämä on luonnollista, sillä näin mitattuna saadut tulonsiirrot tasaavat tuloeroja maksettuja tulonsiirtoja voimakkaammin. Lopputulema on varsin selvä: julkisen sektorin tulontasaustoimien vaikutus kasvoi siirryttäessä vuodesta 1985 vuoteen 1995 ja on tämän jälkeen pienentynyt. Syinä tähän ovat tulonjakotoimien muutokset *sekä* muutokset tuotannontekijätulojen jakaumassa.

6.3 Tilastollinen analyysi

Liitteessä C on esitetty laajennetun Gini-kertoimen arvot käytettävissä oleville tuloille sekä muille tutkituille tulokäsitteille neljää (1,25, 2, 4 ja 10) parametrin ν arvoa käyttäen (taulukko C.1). Aineistoa käsiteltiin siten, että negatiiviset havainnot muutettiin nolla-havainnoiksi ja ekvivalenssiskaalana käytettiin OECD-skaalaa. Mitan arvoille on laskettu bootstrap-keskivirhe-estimaatit bootstrap-otosten määrän ollessa 200. Otosvirheen lisäksi mitan arvoissa on myös ekvivalenssiskaalan valinnasta aiheutunutta virhettä. Virheen suuruutta arvoitiin laskemalla laajennetun Gini-kertoimen arvot myös modifioitua OECD- ja Atkinsonin neliöjuuri-skaalaa käytettäen. Laajennetun Gini-kertoimen arvot poikkesivat toisistaan tyypillisesti alle kaksi keskivirhettä (95 %:n luottamusvälin approksimaatio), mutta joissain tapauksissa erot olivat huomattavia. Siten keskivirhe-estimaatti on hyväksyttävä vain osittaiseksi epävarmuuden mitaksi.

Edellä kuvatuista ongelmista huolimatta, voidaan Gini-kertoimessa tapahtuneiden muutosten tilastollista merkittävyyttä testata tyypillisellä riippumattomien otosten välisellä testillä. Testisuure muodostetaan jakamalla Gini-kertoimien erotus tämän erotuksen keskivirheellä eli

$$Z_{t_1 \to t_2} = \frac{EG(y_{t_2};\nu) - EG(y_{t_1};\nu)}{\sqrt{varEG(y_{t_1};\nu) + varEG(y_{t_2};\nu)}}. \tag{101}$$

Testisuure noudattaa nollahypoteesin ($H_0 : EG(y_{t_1};\nu) = EG(y_{t_2};\nu)$) voimassa olleessa asymptoottisesti standardinormaalijakaumaa.

Liitteen C taulukossa C.2 on esitetty testisuureen $Z_{t_1 \to t_2}$ arvot sekä tähän liittyvän kaksisuuntaisen testin p-arvot (standardinormaalijakauma-approksimaatio). Luvut ilmaisevat selkeästi kehityksen päälinjat. Tuotannontekijätuloissa tuloerot ovat kasvaneet selvästi vuodesta 1985 vuoteen 1995. Tämän jälkeen painotettaessa matalatuloisia, havaitaan tuloerojen kaventumista. Muuten tuloeroissa ei ole ollut merkittävää muutosta. Bruttotuloissa ja käytettävissä olevissa tuloissa tuloerot ovat kasvaneet melko tasaisesti. Kasvun maksimi näyttäisi osuvan periodille 1995 $\rightarrow$ 1998. Painotettaessa matalatuloisia ovat tuloerot vuosina 1985 $\rightarrow$ 1995 pysyneet liki muuttumattomina. Tämän jälkeen kasvu on ollut selvää. Korkeatuloisia painotettaessa tuloerot ovat kasvaneet tasaisesti mutta heikommin.

6.4 Välilliset verot ja verojen kokonaistaakan kohdentuminen

Edellä tarkasteltiin välittömien tulonsiirtojen vaikutusta tuloeroihin. Kotitaloudet maksavat näiden lisäksi kulutuksen kautta ns. välillisiä veroja. Välilliset verot IDT määräytyvät kulutusmenojen määrän ja jakautumisen mukaan. Kun kotitalouden kulutusmenot hyödykeluokassa i ovat c_i ja luokan keskimääräinen veroprosentti on t_i, saadaan

$$IDT(c) = \sum_{j=1}^{K} t_i c_i = \bar{c} \sum_{j=1}^{K} t_i + \sum_{j=1}^{K} t_i(c_i - \bar{c}) = \bar{t} c_{tot} + K cov(t, c), \tag{102}$$

jossa K on hyödykeluokkien lukumäärä ja c_{tot} kotitalouden kokonaiskulutusmenot.

Tarkasteltaessa maksettujen (välillisten tai välittömien) verojen T suhdetta tuloihin y

$$t(y) = \frac{T}{y}, \tag{103}$$

puhutaan keskimääräisestä veroasteesta. Jos keskimääräinen veroaste on veroa edeltävien tulojen kasvava funktio, on verotus progressiivista ja jos se on vähenevä, on verotus regressiivistä (esim. Tuomala 1997, 156). Näin määriteltynä välittömät verot ovat Suomessa progressiivisia ja välilliset verot regressiivisiä. Tämä on helppo todeta analyyttisesti, sillä välillisten verojen asteeksi saadaan

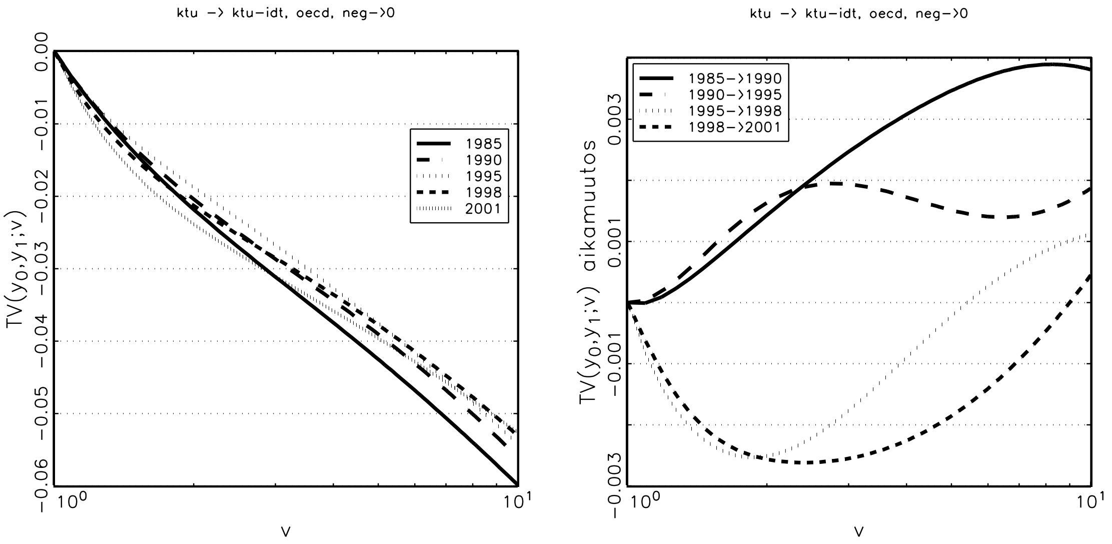
$$t(y) = \bar{t}\,\frac{c_{tot}}{y} + \frac{K cov(t, c)}{y}, \tag{104}$$

jossa tyypillisesti kulutusaste c/y pienenee tulojen kasvaessa. Siten myös veroaste pienenee tulojen kasvaessa, eli välilliset verot ovat regressiivisiä. Pienituloiset siis maksavat enemmän välillisiä veroja tuloihin suhteutettuna kuin suurituloiset. Se kuinka muut tekijät eli keskimääräinen veroprosentti $\bar{t}$ ja kulutuksen ja veroprosentin kovarianssi riippuvat tuloista, on *a priori* epäselvää. Ne voivat joko voimistaa tai heikentää kulutusasteen regressiivistä päävaikutusta. Välillisten verojen regressiivisyys on eräs peruste välittömien verojen progressiivisuudelle.

Empiirisesti tämä voidaan todentaa laskemalla välillisten verojen asteet käytettävissä olevien tulojen desiileissä, mutta myös käyttämällä laajennettuihin Gini-kertoimiin perustuvaa tarkastelua. Tarkasteluja varten määritellään laskennallinen veropuhdistettu tulo, joka saadaan vähentämällä välilliset verot nettotuloista. Tulo on laskennallinen, koska tällaista tulokomponenttia ei todellisuudessa havaita. Välillisten verojen laskemiseksi tarvitaan yhtälön (102) mukaisesti kulutusmenot sekä veroprosentit eri kulutusluokissa i. Kulutusmenot on luokiteltu COICOP-luokissa ja välilliset veroprosentit näille luokille on laskettu panos-tuotos-taulukoiden avulla (Jäntti 2004) ja ne on esitetty Liitteessä D.[37]

Kuvassa 21 on esitetty nettotulojen ja veropuhdistettujen tulojen laajennetun Gini-kertoimen erotus ja tämän erotuksen aikamuutokset. Havaitaan, että TV-funktio saa negatiivisia arvoja. Tuloerot siis kasvavat siirryttäessä nettotuloista veropuhdistettuihin tuloihin. Tämä tarkoittaa sitä, että pienituloiset maksavat enemmän välillisiä veroja suhteessa nettotuloihinsa kuin suurituloiset. Näin ollen välilliset verot ovat regressiivisiä.

[37] Veroprosentit on saatu professori Markus Jäntiltä ja ne on laskenut Sami Toivola Tilastokeskuksessa. On syytä huomata, että panos-tuotos-taulukon hyödykekorit poikkeavat kunkin kotitalouden hyödykekorista hyödykeluokan sisällä. Koska luokkaan kuuluvilla hyödykkeillä on eri veroprosentteja, voivat kotitalouden hyödykeluokan keskimääräiset veroasteet t_i poiketa pano-tuotos-taulokoiden vastaavista. Tämä on huomioitava tuloksia arvioitaessa.

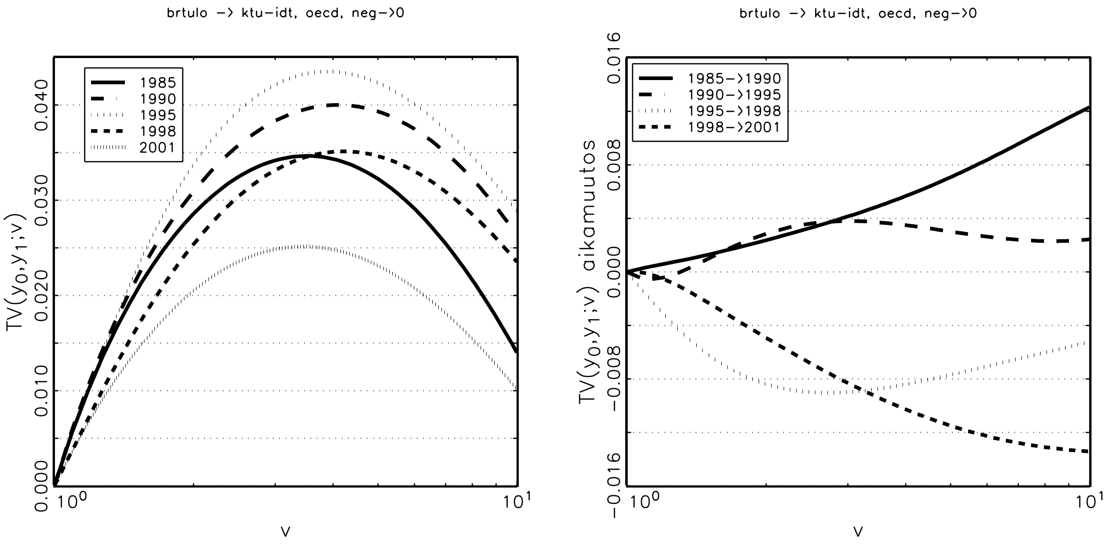


Kuvio 21 Välillisten verojen vaikutus. TV-funktion arvot ovat käytettävissä olevien tulojen ja veropuhdistetun tulon Gini-kertoimien erotus.

Kun kuvaa tarkastellaan lähemmin, havaitaan, että välillisten verojen vaikutusten muutokset ovat moninaisia. Siirryttäessä vuodesta 1985 vuoteen 1995 välillisten verojen regressiivisyys heikkeni kaikilla painotuksilla. Tämän jälkeen (1995 $\rightarrow$ 2001) regressio on kasvanut, paitsi painotettaessa muutoksia pienituloisissa kotitalouksissa (suuret parametrin arvot). Tämän seurauksena painotettaessa suuria tuloja, on välillisten verojen regressio kasvanut ja pienituloisia painotettaessa vähentynyt välillä 1985–2001. Muutokset ovat kuitenkin hyvin pieniä.

Tarkastelu on tähän asti ollut varsin suoraviivaista, mutta jäljellä on vaikea kysymys: mikä on välillisten verojen regressiivisyyden vaikutus hyvinvointiin ja eriarvoisuuteen. Tuloeroista ei tässä yhteydessä ole mielekästä puhua, sillä veropuhdistettu tulo on vain laskennallinen apuväline. Edellä tehdylle tarkastelulle ei siis voida antaa aiempia tarkasteluja vastaavia tuloerotulkintoja. Ilmeistä kuitenkin on se, että jos kotitalouksien maksamat välilliset verot perittäisiin välittöminä tuloista, olisi uuden nettotulojakauman tuloerot vastaavat kuin veropuhdistetun tulon, eli tuloerot kasvaisivat. Ongelmana ovat kuitenkin kotitalouksien erilaiset kulutus- ja siten säästämisasteet. Kotitalouden säästämät varat kulutetaan suurelta osin myöhemmin, jolloin välilliset verot kohdistuvat myös niihin. Onkin esitetty, että välillisiä veroja ei pitäisi suhteuttaa tuloihin vaan kulutusmenoihin (esim. Vartia ja Vartia 1981, luku 3). Tällöin välillisten verojen yhtälöstä (104) jää pois kulutusaste ja samalla poistuu välillisten verojen regressiivisyys. Tarkastelujen erona on se, kuinka säästämiseen ja sen käyttämiseen kulutukseen myöhempinä ajanjaksoina suhtaudutaan. Tässä työssä tyydytään tarkastelemaan verotaakan kohtaantoa, eikä sen hyvinvointivaikutuksiin oteta kantaa.

Verotuksen kokonaistaakan jakautumista tarkastellaan yhdistämällä maksetut tulonsiirrot sekä välilliset verot. Kuvassa 22 on esitetty laajennetun Gini-kertoimen erotus siirryttäessä bruttotulojen jakaumasta veropuhdistetun tulon jakaumaan. Havaitaan, että verotuksen kokonaisvaikutus on progressiivinen koko painotusalueella, joskin parametrin annettaessa kasvaa vielä lisää, voisi verotus aivan pienituloisimpien keskuudessa



Kuvio 22 Verojen kokonaisvaikutus. TV-funktion arvot ovat bruttotulojen ja veropuhdistetun tulon Gini-kertoimien erotus.

olla regressiivistä. Tämä on seurausta pienituloisimpien suurista (yli yksi) kulutusasteista.[38] Samalla havaitaan, että välillisten verojen vaikutus on tyypillisellä parametrialueella vain hieman maksettujen tulonsiirtojen vaikutusta vähäisempää. Aikamuutoksia kuitenkin hallitsee maksettujen tulonsiirtojen muutokset, sillä ne ovat huomattavasti välillisten verojen muutoksia suurempia. Painotettaessa pienituloisia ovat aikamuutokset kummallakin verokomponentilla samansuuntaisia, mutta painotettaessa suurituloisia havaitaan 1990-luvulla verokomponenttien välillä eroja progressiivisuuden kehityksessä. Verotuksen kokonaistaakan jakautumista on tarkastellut perusteellisemmin Jäntti (2004).

6.5 Kulutuseriarvoisuus

Pysyväistulohypoteesin mukaan kotitaloudet suunnittelevat kulutuksensa pitkän aikavälin, hitaasti kehittyvän tulotason eli ns. pysyväistulon $y_{PAV}(t)$ tason mukaan. Tällöin kotitalouden tulot $y(t)$ voidaan esittää muodossa

$$y(t) = y_{PAV}(t) + \delta y(t), \tag{105}$$

jossa $\delta y(t)$ on kotitalouden nopeasti vaihtelevat lyhytkestoiset tuloerät. Pitkän aikavälin (elinkaaren) hyödyn määrittää $y_{PAV}(t)$, jos

$$\int_0^T \delta y(t)dt \approx 0, \tag{106}$$

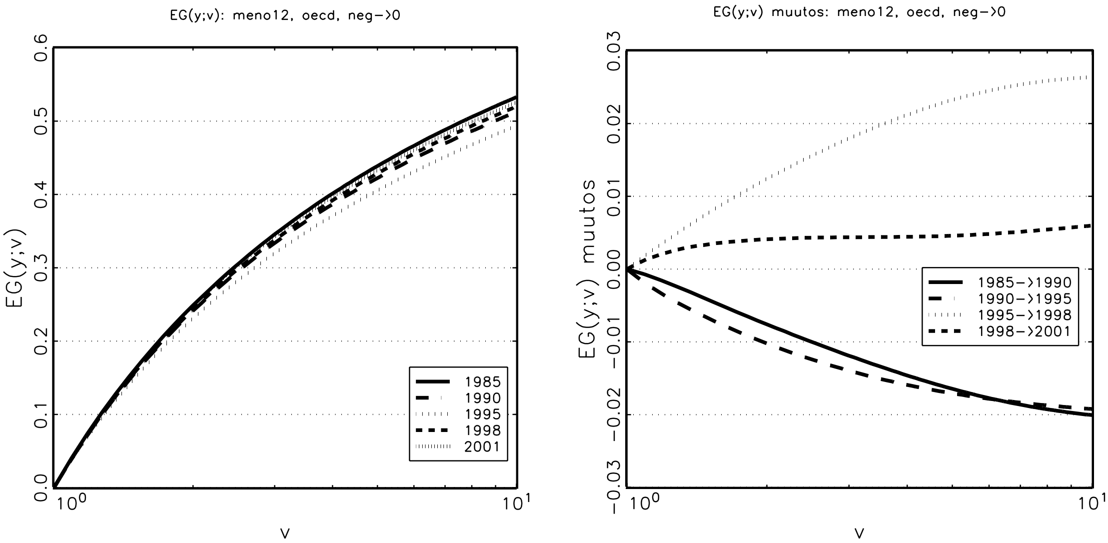
jossa aikaväli $[0, T]$ on pitkä (elinkaari). Koska pysyväistulo määrittää kotitalouksien kulutuksen, voidaan kulutus c muodollisesti esittää pysyväistulon (ja muiden tekijöiden

[38] On syytä huomata, että otosaineistossa on havaintoja, joilla on suuri kulutus, mutta hyvin pienet tulot. Nämä anomaaliset havainnot voivat selittää kokonaisverotuksen mahdollisen regressiivisyyden.

$\mathbf{x}(t))$ funktiona

$$c(t) = C(y_{PAV}(t), \mathbf{x}(t)). \tag{107}$$

Koska taloudellisen hyödyn ajatellaan muodostuvan kulutusvirroista ja koska tulotiedot ovat lyhytaikaisen häiriötulon turmelemat, on kenties perusteltua käyttää kulutusta hyvinvoinnin mittarina (Slesnick 1994). Tällöin tarkastellaan kulutuseriarvoisuutta.



Kuvio 23 Laajennettu Gini-kerroin parametrinsa funktiona ja sen aikamuutokset, kun tulokomponenttina on kotitalouden kulutusmenot. Vertaa kuvaan 11.

Kuvassa 23 on esitetty kulutuseriarvoisuuden tarkastelu laajennettua Gini-kerrointa ja sen ajallisia muutoksia käyttäen. Kuvasta nähdään, että vuosi-aineistolla tarkasteltuna kulutuseriarvoisuus on kehittynyt samansuuntaisesti, mutta kuitenkin eri tavalla kuin tuloeriarvoisuus (kuva 11). Siirryttäessä vuodesta 1985 vuoteen 1995, ovat havaitut kulutuserot laskeneet kaikilla painotuksilla saavuttaen minimin laman syvimmässä vaiheessa (1995). Tämän jälkeen kulutuserot ovat kasvaneet, etenkin taloudellisen nousun kaudella 1995 $\rightarrow$ 1998, mutta kuitenkin niin, että kulutuserot olivat vuonna 2001 hivenen alhaisemmat kuin vuonna 1985. Eroavuudet voivat johtua mm. siitä, että tuloerojen kasvun taustalla on lyhytkestoisen tulon vaihtelun kasvu, joka ei juurikaan vaikuta kulutuksen määrään.

Kuten aiemmin on mainittu, tulojen käyttö hyvinvoinnin mittana kärsii tulojen ajallisesta vaihtelusta johtuvista ongelmista. Etenkin jos suuret tulot ovat myös suuresti vaihtelevia, voivat pitkällä aikavälillä suurituloiset olla vuosiaineistossa hyvinkin pienituloisia. Toisaalta tuloerotarkastelut ovat jo itsessään mielekkäitä, vaikka suoraa hyvinvointivertailua ei voitaisikaan tehdä. Kulutuseriarvoisuuden tarkastelun ongelma on puolestaan siinä, että kaikkea tuloa ei käytetä kulutukseen. Siten on syytä olettaa, että myös tulojen käyttämätön osuus tuottaa saajalleen hyötyä. Muutenhan koko tulo kulutettaisiin aina heti. Pelkän kulutuksen tarkastelu kuvaa vain osaa tulojen kotitaloudelle tuottamasta hyvinvoinnista, eikä siten voi olla yksin riittävä tarkastelun kohde. Voidaankin sanoa, että tulojen ja kulutuksen tarkastelu tukevat toinen toistaan ja niissä havaittavat erot johtavat taustalla olevien tekijöiden syvempään ymmärrykseen (esim. Blundell ja Preston 1998).

6.6 Kansainvälinen vertailu

Jotta Suomessa havaitun eriarvoisuuden määrän voisi paremmin ymmärtää, on arvoja syytä verrata kansainvälisesti. Taulukkoon 3 on kerätty eräiden OECD-maiden Gini-kertoimen arvoja vuodelta 2000 (Förster ja Mira d'Ercole 2005). Taulukossa on myös Suomi, jonka arvot perustuvat lähteen laskelmiin. Näin Suomea voidaan vertailla muihin maihin ilman, että poikkeava vuosi tai eriävä aineisto haittaisi tarkastelua. Taulukosta nähdään, että markkinoiden synnyttämät tuloerot ovat Suomessa pienet ja julkisen sektorin tulontasaustoimet melko voimakkaita. Edellä mainitun tutkimuksen mukaan, vuonna 2000 Suomessa oli OECD-maiden 7. pienin tuloeriarvoisuus Gini-kertoimella mitattuna. OECD-maat näyttäisivät jakaantuvan pienten tuloerojen (Benelux-maat, Pohjoismaat sekä Itävalta ja Tšekki), keskimääräisten tuloerojen (Ranska, Saksa, Unkari, Espanja, Irlanti, Kanada sekä Australia) ja yli keskiarvoisten tuloerojen (Japani, Iso-Britannia, Uusi-Seelanti, Kreikka, Italia, Portugali, Puola ja USA) ryhmiin. Suurimmassa osassa OECD-maita tuloerot (Gini-kerroin) ovat kasvaneet 1980-luvun puolivälistä 1990-luvun puoliväliin. Tämän jälkeen (→ 2000) tuloerot ovat kasvaneet voimakkaimmin Suomessa ja Ruotsissa. Tämä voidaan nähdä eräänlaisena konvergoitumisprosessina, joskin samaan aikaan Ranska, Saksa ja Alankomaat pystyivät hieman vähentämään tuloerojaan. (Förster ja Mira d'Ercole 2005.)

Taulukko 3 Vuoden 2000 Gini-kertoimen arvot (%) työikäisten markkinatuloille (tuotannontekijätulot, TT-tulo) ja käytettävissä oleville tuloille (KTU).

Maa	TT-tulo	KTU	erotus
Alankomaat	33,2	25,1	8,1
Iso-Britannia	43,2	32,6	10,6
Ruotsi	37,5	24,3	13,2
Saksa	39,3	27,7	11,6
Suomi	37,1	26,1	11,0
Yhdysvallat	42,0	35,7	6,3

Lähde: Förster ja Mira d'Ercole (2005).

7 Johtopäätökset

Tutkimuksessa on tarkasteltu tuloerojen ja taloudellisen eriarvoisuuden mittaamisen teoriaa sekä tuloerojen ja taloudellisen eriarvoisuuden muutoksia Suomessa 1985–2001. Tulokset ovat siten jakaantuneet kahteen osaan. Pääpaino on tuloerojen mittaamisen teorialla, etenkin ns. eriarvoisuusmittojen ominaisuuksien ja tilastollisten ominaisuuksien tarkastelulla. Tuloerojen empiirinen kuvaaminen muodostaa tulosten toisen osan. Aineistona käytettiin Tilastokeskuksen Kulutustutkimuksia vuosilta 1985, 1990, 1994–1996, 1998 sekä 2001. Tuloerojen tarkasteluun on Suomessa muitakin aineistoja, kuten esimerkiksi Tulonjakotilasto. Tutkimuksessa kiinnostuksen kohteena on tulojakaumien lisäksi myös kulutusjakauma sekä välillisten verojen kehitys. Kulutusmenotietojen tarpeen vuoksi aineistona käytettiin Kulutustutkimuksia.

7.1 Eriarvoisuuden mittaamisesta ja mitoista

Taloudellisen eriarvoisuuden mittaaminen on vaativa tehtävä ja kirjallisuudessa on esitetty lukuisia teemaan sisältyviä tuloksia (esim. Nygård ja Sandström 1981, Sen 1973). Tehtävään liittyy lukuisia vaikeuksia, joista eräs merkittävä on eriarvoisuuden käsitteen määrittely. Jos määrittelemme eriarvoisuuden tulojen tuottaman taloudellisen hyvinvoinnin hajaantumisena, niin ajattelemme, että yhteiskunnassa tulee oikeudenmukaisuuden nimissä olla tuloeroja ja siten myös näin määriteltyä eriarvoisuutta. Tämä seuraa jo yksin siitä, että vaivannäöstä on oikeudenmukaista saada korvaus ja yksilöt näkevät eri määrän vaivaa tulojensa eteen. Toisaalta voimme nähdä eriarvoisuuden niinä tuloeroina, jotka eivät ole erilaisuuden oikeuttamia. Näitä oikeutettuja tuloeroja on mahdotonta mitata, ennen kuin tuloerot oikeuttava erilaisuus ja oikeutettujen tuloerojen määrä on määritelty. Siten käytettävissä olevilla aineistoilla on tyydyttävä tarkastelemaan ensin kuvattua eriarvoisuutta.

Tulojen tuottaman hyvinvoinnin hajontana määritetty eriarvoisuus johtaa ongelmiin eriarvoisuuden mittaamisen teorian kanssa. Teoreettiset tarkastelut kohdistuvat nimittäin tilanteeseen, jossa yhteiskunnallisesti optimaalista olisi antaa kaikille kotitalouksille saman verran ekvivalenttituloja (esim. Atkinson 1970). Tämä ei varmastikaan olisi oikeudenmukaisin tulonjako. Siten teoria ei pysty antamaan vastausta kysymykseen, mikä on optimaalinen tuloerojen taso. Tämän seurauksena tuloerojen ja eriarvoisuuden mittaaja ei voi tehdä muuta kuin ilmoittaa kehityksen suunnan. Hän ei siis periaatteessa voi tehdä tilanteeseen soveltuvaan teoriaan pohjautuen johtopäätöstä siitä, onko kehitys mennyt yhteiskunnan kannalta hyvään vai huonoon suuntaan.

Eriarvoisuuden erilaisista määritelmistä voi seurata virheellinen ajatus siitä, että kaikki havaittu eriarvoisuus olisi pahasta. Samoin kasvaneet tuloerot koetaan usein lisääntyneenä määränä hyvin pienituloisia sekä hyvin suurituloisia. Näin ei kuitenkaan tarvitse olla, sillä vain toisen ryhmän suhteellinen kasvu riittää. Suhteellisena maksajana tai saajana voivat olla keskituloiset. Eriarvoisuutta mitattaessa on pyrittävä selvittämään, missä kohdassa tulojakaumaa muutokset tapahtuvat. Tällä on merkitystä havaitun muutoksen hyvinvointivaikutusten arvioinnissa. On myös syytä huomata, että kasvava eriarvoisuus voi myös hyvin kuvata yhteiskunnan sallivammaksi muuttuneita asenteita, jolloin tuloerojen kasvun ei katsota vähentävän hyvinvointia. Jos tuloerojen lähtötaso on korkea, on melko turvallista olettaa, että tuloerojen laskua voidaan yhteiskunnan kannalta pitää hyvinvointia lisäävänä ilmiönä. Tuloerojen muutosten tarkka kuvaaminen mahdollistaa voittajien ja häviäjien paikantamisen ja tämän tiedon avulla voidaan paremmin arvioida hyvinvoinnin muutoksia.

Tutkimuksessa keskityttiin taloudelliseen ja suhteelliseen eriarvoisuuteen, joka ajateltiin tulojen tuottaman hyvinvoinnin suhteelliseksi hajonnaksi. Eriarvoisuuden mittaamiseen on kehitetty lukuisia mittaperheitä, joilla on kullakin laaja parametrialueensa, josta valita eriarvoisuuden luonnetta tarkentava parametrin arvo. Siten mitoilla voidaan kvantifioida jakauman muutoksia monilla erilaisilla tulopainotuksilla. Tämän vuoksi tyytyminen yhteen mittaan ja sen yhteen parametrin arvoon ei ole mielekästä, jollei tutkija pysty täsmällisesti perustelemaan valintaansa. Sen sijaan jo kolmella parametrin arvolla saadaan huomattavasti enemmän tietoa jakaumien välisistä muutoksista, kun painotukset tehdään järkevästi. Tutkimuksessa havaittiin esimerkiksi , että toisilla painotuksilla tuotannontekijätuloille laskettu mitan arvo kasvaa, kun toisilla se pienenee (ks. Liite C). Tämä ilmaisee muutoksen sijainnin jakaumassa. Toisaalta mittoja voidaan tarkastella jopa parametrinsa funktiona graafisesti, jolloin tutkija ei ole parametrinvalinnan vanki.

Eriarvoisuusmittojen käytössä on eräitä käytännön vaikeuksia. Eräs vaikeus liittyy siihen, että populaatio on heterogeeninen. Erilaiset kotitaloudet vaativat eri määrän tuloja saadakseen saman hyvinvoinnin tason. Tutkimuksessa ongelma on ratkaistu ekvivalenssiskaaloilla, joiden teoreettinen perustelu nojaa Ebertin (1997 ja 1999) tutkimuksiin.

Toinen käytännön vaikeus liittyy siihen, että vuositason tulot eivät täydellisesti vastaa kotitalouden taloudellisen hyvinvoinnin tasoa. Tämä johtuu siitä, että tulot vaihtelevat mahdollisesti suurestikin vuodesta toiseen. Ongelma on tässä työssä ratkaistu olettamalla, että vuosittainen vaihtelu ei ole merkittävästi muuttunut tutkitulla periodilla, ja siten tulojakauman muutokset antavat kohtuullisen kuvan myös hyvinvoinnin jakauman muutoksista. Lisäinformaationa käytetään tutkimuksessa kotitalouksien kulutusmenoja, joiden ajatellaan kuvaavan kotitalouden sen hetkistä hyvinvointia. Voidaan kuitenkin perustellusti ajatella, että tulot tuottavat muutakin hyvinvointia kuin vain sen hetken kulutuksen, ja siten tulo- ja kulutustietoja voidaan käyttää tukemaan toisiaan.

Kolmas käytännön vaikeus liittyy tuloerojen ja eriarvoisuuden suhteeseen. Eriarvoisuusmitat pyrkivät nimensä mukaisesti kuvaamaan pääasiassa eriarvoisuutta ja tämä johtaa siihen, ettei tuloerojen syntyä aina pystytä tutkimaan taustalla olevien tulokomponenttien avulla. Tämä johtuu siitä, että eriarvoisuuden mittaamisen teoria on tyypillisesti kehitetty jakaumille, joissa tulot ovat positiivisia. Siten myös mitat on kehitetty käytettäväksi tällaisiin jakaumiin. Käytettävissä olevat tulot, joille on mielekkäintä antaa eriarvoisuustulkinta, koostuvat monista tulokomponenteista, joiden tuloerojen muutoksia olisi myös mielekästä tarkkailla, esimerkiksi tutkittaessa julkisen sektorin tulontasaustoimien muutoksia. Tämä ei kuitenkaan aina vuositason aineistolla onnistu, sillä eräillä tulokomponenteilla on huomattavasti negatiivisia arvoja. Siten mittojen voimakas keskittyminen eriarvoisuuteen, tuloerojen mittaamisen sijasta, vähentää näiden mittojen käytettävyyttä.

Tutkimuksessa tarkasteltujen mittojen välillä havaittiin merkittäviä eroja. Teoreettiselta taustaltaan ne edustavat kolmea erilaista lähestymistapaa eriarvoisuuden mittaamisessa: laajennettu Gini-kerroin on sopiva tilastollinen tunnusluku, Atkinsonin mitta on yhteiskunnan preferenssejä painottava mitta ja yleistetty entropia on eriarvoisuuden mittaamisen teoriassa yleistetty informaatioteorian käsite. Käytännön mittaustyössä näillä taustojen eroilla ei ole merkitystä, vaan tärkeintä on mitan käyttökelpoisuus ja mitan mahdollistamat erilaiset painotukset.

Tässä työssä sopivimmaksi mitaksi osoittautui laajennettu Gini-kerroin lähinnä muiden mittojen kärsimien ongelmien, ei niinkään Gini-kertoimen erinomaisuuden vuoksi. Laajennetun Gini-kertoimen etuna on se, että se sietää tarvittaessa negatiivisia ja nollahavaintoja, eikä se kärsi suuresti poikkeavista havainnoista. Kaksi muuta esiteltyä mittaa on näiltä osin puutteellisia. Mutta kuten luvussa 3 havaittiin, ei laajennettu Gini-kerroin pysty kuvaamaan jakaumaa samoin painotuksin kuin nämä kaksi muuta mittaa. Siten Atkinsonin indeksin ja yleistetyn entropian käytölle on hyvät perustelut, kunhan niiden rajoitukset huomioidaan. Näitä mittoja voisi kenties hyödyntää bootstrap-tyyppisellä analyysillä, jossa poikkeavien havaintojen vaikutus tulisi selvemmin kontrolloitua. Tämän lisäksi on syytä huomioida se, että yleisesti käytetty tavallinen Gini-kerroin on eriarvoisuusmittana ongelmallinen. Gini-kerroin ei reagoi herkästi hyvin pienien tulojen kasvuun. Siten voitaneen sanoa, että Gini-kerroin ei ole erityisen hyvä eriarvoisuusmittari, jos mittarilla halutaan tuoda esiin populaatiossa esiintyvät hyvin pienituloiset taloudet.

7.2 Tuloerojen kehitys Suomessa

Empiirisessä osiossa havaittiin, että tuloeriarvoisuus laajennetulla Gini-kertoimella mitattuna on kasvanut vuodesta 1985 vuoteen 2001 siirryttäessä. Mielenkiintoinen piirre tuloksissa on se, että tuotannontekijätulojen tuloerot kasvoivat välillä 1990–1995, mutta käytettävissä olevat tulojen vasta välillä 1995–2001. Tämä merkitsee sitä, että tuotannontekijätulojen rakenteessa sekä julkisen sektorin toimissa on tapahtunut muutos, joka on johtanut julkisen sektorin tuloja tasaavan vaikutuksen heikkenemiseen vuosituhannen lopulla.

Tuotannontekijätulojen tuloerojen kasvu on laman jälkeen pysähtynyt ja jopa kääntynyt laskuun, kun painotetaan matalia tuloja. Syynä tuloerojen laskuun lienee laman jälkeen hitaasti parantuva työllisyystilanne. Muutokset ovat laman jälkeen olleet tilastollisesti merkitseviä vain painotettaessa kaikkein pienituloisimpia. Käytettävissä olevien tulojen tapauksessa tuloerot ovat kasvaneet laman jälkeen kaikilla painotuksilla[39]. Kasvu on tilastollisesti merkitsevintä painotettaessa pienituloisia. Myös kulutuksessa on havaittu vuosituhannen lopulla eriarvoisuuden kasvua, mutta kulutuseriarvoisuus laski laman aikana niin alhaiseksi, ettei se vuonna 2001 ollut yhtä suurta kuin vuonna 1985.

Tulonsiirtojen ja välillisten verojen vaikutusta tarkasteltiin eri tulokäsitteiden laajennetun Gini-kertoimen arvojen erotuksina. Tunnusluvuille ei tehty tilastollisia merkit-

[39] Käytettävissä olevien tulojen hajonnalle on tutkituista tulokäsitteistä mielekkäintä antaa eriarvoisuustulkinta

sevyystarkasteluja. Saatujen tulonsiirtojen kehitys on ollut kaksijakoinen. Periodilla 1985–1995 saatujen tulonsiirtojen tulonjakoa tasaava vaikutus voimistui ja välillä 1995–2001 se heikkeni kaikilla painotuksilla. Maksettujen tulonsiirtojen eli lähinnä välittömien verojen tulonjakovaikutusten kehitys ei ole ollut yhtä suoraviivaista. Jaksolla 1985–1990 maksettujen tulonsiirtojen tulonjakoa tasaava vaikutus (progressio) kasvoi painotettaessa pienituloisia. Muuten muutokset olivat välillä 1985–1995 pieniä ilman selvää suuntaa. Vuosina 1995–2001 tulonjakoa tasaava vaikutus heikkeni painotettaessa pienituloisia. Suurituloisia painotettaessa heikentävä muutos ajoittuu välille 1995–1998. Tulonsiirtojen yhteisvaikutus on saatujen tulonsiirtojen dominoimaa: Periodilla 1985–1995 saatujen tulonsiirtojen tulonjakoa tasaava vaikutus voimistui ja välillä 1995–2001 se heikkeni kaikilla painotuksilla. Tulonsiirtojen tulontasausvaikutusten muutokset johtuvat tuotannontekijätulojen tuloerojen ja koostumuksen muutoksista sekä julkisen vallan toimien muutoksista. On syytä huomata, että tuotannontekijätuloissa tuloerojen muutos on ollut melko vähäistä vuosina 1995–2001, joten tulonjakovaikutusten muutokset tällä periodilla johtunevat pääsääntöisesti tulorakenteen ja lainsäädännön muutoksista.

Välillisten verojen regressiivisyyden muutokset ovat olleet melko vähäisiä. Periodilla 1985–1995 regressiivisyys on vähentynyt, mutta tämän jälkeen se on kääntynyt kasvuun, kun ei painoteta pienituloisia. Painotettaessa pienituloisia periodilla 1995–2001 muutokset ovat hyvin pieniä. Välillisiä veroja koskevat tulokset ovat lähinnä suuntaa-antavia, sillä käytetty hyödykeluokittelu on melko karkea.

Kansainvälisessä vertailussa Suomen tuloerojen kehitys poikkeaa hieman muista OECD-maista. Tuloerot ovat kasvaneet suurimmassa osassa OECD-maita välillä 1985–1995 ja tämän jälkeen kehitys on tasoittunut. Sen sijaan Suomessa ja Ruotsissa tuloerojen kasvu jatkunut voimistuen välillä 1995–2000. Tuloerot ovat Suomessa kuitenkin kansainvälisesti vertaillen melko alhaiset. (Förster ja Mira d'Ercole 2005.)

7.3 Jatkossa tehtävät tarkastelut

Tässä työssä käytettiin eräitä oikopolkuja ja jätettiin tekemättä tarkasteluja, jotka jatkotutkimuksissa olisi kenties syytä tehdä. Eräs oikopolku on OECD-ekvivalenssiskaalan käyttö. Vaikka herkkyystarkastelut osoittivat, että skaalan käyttö vaikuttaa harmittomalta yksinkertaistukselta, olisi silti paikallaan pyrkiä estimoimaan aineistolle omat, paremmin aineistoa vastaavat ekvivalenssiskaalat. Vakioisen ekvivalenssiskaalan käyttö voi nimittäin tuoda esiin kuvitteellisia efektejä, jotka johtuvat kotitalousrakenteen ja tulojen hyvinvointivaikutusten välisen suhteen muutoksista vuosien saatossa. Tutkittaessa näin lyhyttä aikaperiodia (1985 $\rightarrow$ 2001) tästä ei kuitenkaan aiheutune suurta virhettä. Sen sijaan rakenteen jatkuva väärin huomioiminen voi tuottaa virheellisiä johtopäätöksiä. Ekvivalenssiskaalojen estimoiminen vaatisi luonnollisesti oman tutkimuksen.

Oikopolkuna voidaan kenties pitää myös sitä, että tutkimuksessa käytettiin bootstrap-menetelmää keskivirheiden estimointiin. Mielenkiintoista olisi verrata saatuja estimaatteja asymptoottisen teorian tuloksiin ja esimerkiksi Monte Carlo -simulaatioin tarkastella menetelmien ominaisuuksia.

Lisätietoa tuloerojen kehityksestä saataisiin tutkimalla kokonaisuutta osiensa kautta. Eräs tapa olisi tarkastella tuloerojen kehittymistä taustamuuttujien määrittelemissä osaryhmissä. Näin saataisiin tietoa siitä, keitä muutokset lähinnä koskevat ja mahdollisesti myös siitä mistä muutokset johtuvat. Toinen tapa tarkentaa tutkimusta olisi mitata tuloerojen muutoksia tuotannontekijätulon eri komponenteissa. Näin saataisiin tietoa, onko tuloerojen muutosten taustalla esimerkiksi palkka- vai omaisuustuloissa tapahtuneet muutokset. Pitäisi myös pohtia kulutus- ja tulotietojen käyttämistä yhdessä, jotta pysyväistulohypoteesin mukainen tulon hyvinvointia tuottava pysyvä komponentti voitaisiin paremmin määritellä.

Lähteet

Ahlqvist, K. ja Pajunen, A. (2000): *Kotitalouksien kulutusmenojen muutokset 1990-luvulla, Tulot ja kulutus 2000:28.* Yliopistopaino, Helsinki.

Arponen, J. (1994): *Statistinen fysiikka.* 2. painos, Limes, Helsinki.

Atkinson, A. B. (1970): On the measurement on inequality. *Journal of Economic Theory*, 2, 244–263.

Atkinson, A. B. ja Bourguignon, F. (1987): Income distribution and and differences in needs. Teoksessa Feiwel, G. R. (toim.) *Arrow and the Foundations of the Theory of Economic Policy*, luku 12, Macmillan, London.

Aura, S. (1996): *Lorenz-käyrät, hyvinvointiteoriat ja tilastollinen päättely, VATT-tutkimuksia*, osa 35. VATT, Helsinki.

Barrett, G. F. ja Pendakur, K. (1995): The asymptotic distribution of the generalized Gini indices of inequality. *The Canadian Journal of Economics*, 28(4b), 1042–1055.

Beach, C. M. ja Davidson, R. (1983): Distribution-free statistical inference with Lorenz curves and income shares. *Review of Economic Studies*, 50(4), 723–735.

Blundell, R. ja Preston, I. (1998): Consumption inequality and income uncertainity. *The Quarterly Journal of Economics*, 113(2), 603–640.

Brilinger, D. R. (1966): Discussion on mr Sprent's paper. *Journal of the Royal Statistical Society, Series B*, 28(2), 294.

Chotikapanich, D. ja Griffiths, W. (2001): On calculation of the extended Gini coefficient. *The Review of Income and Wealth*, 47(4), 541–548.

Cowell, F. A. (1989): Sampling variance and decomposable inequality measures. *Journal of Econometrics*, 42, 27–41.

——— (1995): *Measuring Inequality.* LSE Handbooks in Economics, 2. painos, Prentice Hall, London.

——— (2000): Measurement of inequality. Teoksessa Atkinson, A. B. ja Bourguignon, F. (toim.) *Handbook of income distribution*, *Handbook in economics*, osa 16, 87–166, Elsevier, Amsterdam.

Cowell, F. A. ja Kuga, K. (1981): Additivity and the entropy concept: An axiomatic approach to inequality measurement. *Journal of Economic Theory*, 25(1).

Dalton, H. (1920): The measurement of the inequality of incomes. *The Economic Journal*, 30(119), 348–361.

Dasgupta, P., Sen, A. ja Starrett, D. (1973): Notes on the measurement of inequality. *Journal of Economic Theory*, 6(2), 180–187.

Deltas, G. (2003): The small-sample bias of the Gini coefficient: results and implications for empirical research. *The Review of Economics and Statistics*, 85(1), 226–234.

Donaldson, D. ja Weymark, J. A. (1980): A single parameter generalization of the Gini indices pf inequality. *Journal of Economic Theory*, 22(1), 67–86.

Dorfman, R. (1979): A formula for the Gini coefficient. *The Review of Economics and Statistics*, 61(1), 146–149.

Ebert, U. (1997): Social welfare when needs differ: An axiomatic approach. *Economica*, 64(254), 233–244.

——— (1999): Using equivalent income of equivalent adults to rank income distributions. *Social Choice and Welfare*, 16(2), 233–258.

Efron, B. (1979): Bootstrap methods: another look at the jackknife. *The Annals of Statistics*, 7(1), 1–26.

——— (1981): Nonparametric estimates of standard error: the jackknife, the bootstrap and other methods. *Biometrika*, 68(3), 589–599.

Efron, B. ja Gong, G. (1983): A leisurely look at the bootstrap, the jackknife and cross-validation. *The American Statistician*, 37(1), 36–48.

Efron, B. ja Tibshirani, R. (1986): Bootstrap methods for standard errors, confidence intervals and other measures of statistical accuracy. *Statistical Science*, 1(1), 54–75.

Efron, B. ja Tibshirani, R. J. (1993): *An Introduction to the Bootstrap*, *Monographs on Statistics and Applied Probability*, osa 57. Chapman & Hall, New York.

Förster, M. ja Mira d'Ercole, M. (2005): Income distribution and poverty in OECD countries in the second half of the 1990s. *OECD Social and Migration Working Papers*, 22.

Gastwirth, J. L. (1971): A general definition of the lorenz curve. *Econometrica*, 39(6), 1037–1039.

Gini, C. (1912): *Variabilità e mutabilità*, *Studi economico-giuridici della r. universita di Cagliari*, osa 3.

——— (1914): Sulla misura della concentrazione e della variabilità dei caratteri. *Atti del R. Istituto Veneto di Scienze, Lettere e Arti*, 73, 1203–1248.

——— (1921): Measurement of inequality of incomes. *The Economic Journal*, 31(121), 124–126.

Hagfors, R. (1988): *Kotitalouksien ekvivalenssiskaalat Suomessa; Empiirinen poikkileikkaustarkastelu*, *Sarja C*, osa 46. ETLA, Helsinki.

Jäntti, M. (2004): The distribution of the tax burden in Finland, http://www.abo.fi/˜mjantti/TaxesandIncomeDistribution.pdf.

Kahneman, D. ja Tversky, A. (1979): Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–292.

Kakwani, N. (1980a): *Income Inequality and Poverty; Methods of Estimation and Policy Applications*. A World Bank Research Publication, Oxford University Press, New York.

——— (1980b): On a class of poverty measures. *Econometrica*, 48(2), 437–446.

Lambert, P. J. (1985): Social welfare and the Gini coefficient revisited. *Mathematical Social Sciences*, 9(1), 19–26.

——— (2001): *The Distirbution and Redistribution of Income*. 3. painos, Manchester University Press, Manchester.

Lerman, R. I. ja Yitzhaki, S. (1984): A note on the calculation and interpretation of the Gini index. *Economics Letters*, 15(3–4), 363–368.

——— (1985): Income inequality effects by income source: a new approach and applications to the united States. *the Review of Economics and Statistics*, 67(1), 151–156.

——— (1995): Changing ranks and the inequality impacts of taxes and transfers. *National Tax Journal*, 48(1), 45–59.

Mehran, F. (1976): Linear measures of income inequality. *Econometrica*, 44(4), 805–809.

Miller, R., Jr. (1974): The jackknife – a review. *Biometrika*, 61(1), 1–15.

Moyes, P. (1987): A new concept of lorenz domination. *Economics Letters*, 23(2), 203–207.

Musgrave, R. A. ja Thin, T. (1948): Income tax progression, 1929–48. *The Journal of Political Economy*, 56(6), 498–514.

Nygård, F. ja Sandström, A. (1981): *Measuring income inequality*, *Stockholm Studies in Statistics*, osa 1. Almqvist & Wiksell International, Stockholm.

Quenouille, M. H. (1949): Approximate tests of correlation in time-series. *Journal of the Royal Statistical Society. Series B*, 11(1), 68–84.

——— (1956): Notes on bias estimation. *Biometrika*, 43(3/4), 353–360.

Riihelä, M. ja Sullström, R. (2002): Käytettävissä olevien tulojen liikkuvuus Suomessa 1990–1999. *VATT keskustelualoitteita*, 270.

——— (2004): Välittömien verojen ja tulonsiirtojen vaikutus tulonsaajajärjestykseen ja tuloerojen muutoksiin Suomessa. *VATT keskustelualoitteita*, 345.

Riihelä, M., Sullström, R. ja Tuomala, M. (2005): Trends in top income shares in Finland. *VATT keskustelualoitteita*, 371.

Sen, A. K. (1973): *On Economic Inequality*. Oxford University Press, London.

Shorrocks, A. F. (1983): Ranking income distributions. *Economica*, 50(197), 3–17.

Slesnick, D. T. (1994): Consumption, needs and inequality. *International Economic Review*, 35(3), 677–703.

Suoniemi, I. ja Sullström, R. (1995): *The structure of household consumption in Finland, 1966–1990*, *VATT-tutkimuksia*, osa 27. VATT, Helsinki.

Theil, H. (1967): *Economics and Information Theory*. North Holland, Amsterdam.

Thistle, P. D. (1990): Large sample properties of two inequality indices. *Econometrica*, 58(3), 725–728.

Tilastokeskus (2001): *Kulutustutkimus 1998 laatuselvitys, Tulot ja kulutus 2001:4*. Yliopistopaino, Helsinki.

——— (2002): *Yksilöllisen kulutuksen käyttötarkoituksen mukainen luokitus (COICOP)*, *Käsikirjoja*, osa 44. Tilastokeskus, Helsinki.

Tukey, J. W. (1958): Bias and confidence in not-quite large samples. *The Annals of Mathematical Statistics*, 29(2), 614–623.

Tuomala, M. (1997): *Julkistalous*. Gaudeamus, Tampere.

Vartia, P. L. I. ja Vartia, Y. O. (1981): *Liikevaihtoveron korottaminen ja tulojen ostovoima*, *Sarja B*, osa 31. ETLA, Helsinki.

Vartia, Y. (1988): Ovatko tilastolliset tutkimukset vain monimutkaisesti perusteltuja mielipiteitä. *Kansantaloudellinen aikakauskirja*, 84(3), 305–309.

——— (1989): *Tilastotieteen perusteet*. 2. painos, Yliopistopaino, Helsinki.

Wolter, K. M. (1985): *Introduction to Variance Estimation*. Springer-Verlag, New York.

Yitzhaki, S. (1983): On a extension of the Gini coefficient. *International Economic Review*, 24(3), 617–628.

Liite A: Välivaiheita lukuun 3

Jaksoon 3.2.1 liittyvät välivaiheet

Gini-kertoimen määritelmän mukainen (yhtälö (24)) integraali on helppo laskea, kun huomataan, että Lorenz-käyrän määritelmä jatkuvassa muodossa (jakso 2.3.1) johtaa siihen, että Lorenz-käyrän rajaama ala muodostuu n kappaleesta puolisuunnikkaita, joiden ala on $w_i(\Phi_i + \Phi_{i-1})/2$, $i \in \{1, 2, \ldots, n\}$, jossa $\Phi_0 = 0$ ja $\Phi_0 = 1$. Täten Gini-kertoimeksi saadaan

$$G = 1 - \sum_{j=1}^{n} w_j(\Phi_{j-1} + \Phi_j).$$

Sijoittaen edelliseen yhtälöön Lorenz-ordinaatan määritelmä

$$\Phi_i = \sum_{j=1}^{i} w_j \frac{y_j}{\overline{y}},$$

saadaan

$$G = 1 - \frac{2}{\overline{y}} \left[\sum_{j=1}^{n} \sum_{k=1}^{j} w_j w_k y_k - \frac{1}{2} \sum_{j=1}^{n} w_j^2 y_j \right].$$

Kaksoissummaa tutkimalla havaitaan, että

$$\sum_{j=1}^{n} \sum_{k=1}^{j} w_j w_k y_k = w_n^2 y_n + y_{n-1} w_{n-1}(w_{n-1} + w_n) + \cdots + y_2 w_2(1 - w_1) + y_1 w_1,$$

eli

$$\sum_{j=1}^{n} \sum_{k=1}^{j} w_j w_k y_k = \sum_{j=1}^{n} (p_n - p_{j-1}) w_j y_j = \overline{y} - \sum_{j=1}^{n} p_{j-1} w_j y_j,$$

jossa $p_i = \sum_{j=1}^{i} w_j$ ja $p_0 = 0$ sekä $p_n = 1$. Sijoittamalla saatu muoto Gini-kertoimen lausekkeeseen, päädytään yhtälöön

$$G = \frac{2}{\overline{y}} \sum_{j=1}^{n} (p_{j-1} w_j y_j + \frac{1}{2} w_j^2 y_j) - 1.$$

Lisäämällä ja vähentämällä sulkeiden sisälle $w_j^2 y_j$ saadaan lopullinen muoto

$$G = \frac{2}{\overline{y}} \sum_{j=1}^{n} (p_j w_j y_j - \frac{1}{2} w_j^2 y_j) - 1.$$

Tähän muotoon päädytään myös, jos Gini-kertoimen integraalimääritelmän integrointi suoritetaan vaakatasossa olevien puolisuunnikkaiden avulla tai jos lähtökohdaksi otetaan yhtälö (30) ja poistetaan itseisarvot käyttämällä relaatiota

$$|y_j - y_k| = y_j + y_k - 2\min(y_j, y_k).$$

Jaksoon 3.2.2 liittyvät välivaiheet

Laajennettu Gini-kerroin on määritelty yhtälöllä (35), joka voidaan jakaa integrointeihin Lorenz-käyrän segmenttien yli eli

$$EG(\nu) = 1 - \nu(\nu-1)\sum_{j=1}^{n}\int_{p_{j-1}}^{p_j} L(p)(1-p)^{\nu-2}\mathrm{d}p.$$

Prosenttipisteiden p_{i-1} ja p_i välillä Lorenz-käyrä on määritelty yhtälöllä

$$\begin{aligned} L(p) &= L_{i-1} + \frac{y_i}{\overline{y}}(p - p_{i-1}) \\ &= A_i + r_i p, \end{aligned}$$

kun $A_i = L_{i-1} - r_i p_{i-1}$ ja $r_i = y_i/\overline{y}$. Näin mielenkiinnon kohteena olevat integraalit voidaan kirjoittaa muodossa

$$I_i = \int_{p_{i-1}}^{p_i} (A_i + r_i p)(1-p)^{\nu-2}\mathrm{d}p$$

ja siten jakaa kahteen erilliseen integraaliin

$$I_i = A_i\int_{p_{i-1}}^{p_i}(1-p)^{\nu-2}\mathrm{d}p + r_i\int_{p_{i-1}}^{p_i} p\,(1-p)^{\nu-2}\mathrm{d}p.$$

Näistä ensimmäinen saadaan laskettua suoraan integroimalla

$$\int_{p_{i-1}}^{p_i}(1-p)^{\nu-2}\mathrm{d}p = -\frac{1}{\nu-1}\Big/_{p_{i-1}}^{p_i}(1-p)^{\nu-1}$$

ja toinen osittaisintegroimalla

$$\int_{p_{i-1}}^{p_i} p\,(1-p)^{\nu-2}\mathrm{d}p = -\frac{1}{\nu-1}\Big/_{p_{i-1}}^{p_i}\left[p\,(1-p)^{\nu-1} + \frac{1}{\nu}(1-p)^{\nu}\right].$$

Tulokset yhdistämällä saadaan integraalille arvoksi

$$I_i = -\frac{1}{\nu-1}\Big/_{p_{i-1}}^{p_i}\left[(A_i + r_i p)(1-p)^{\nu-1} + \frac{r_i}{\nu}(1-p)^{\nu}\right].$$

Integraali on nyt laskettu, mutta tulosta voidaan vielä sieventää. Tarkastellaan lähemmin sijoituslausekkeen ensimmäistä termiä sijoittamalla siihen A_i:n määritelmä

$$\begin{aligned} \Big/_{p_{i-1}}^{p_i}(A_i + r_i p)(1-p)^{\nu-1} &= (L_{i-1} + r_i\Delta p_i)(1-p_i)^{\nu-1} - L_{i-1}(1-p_{i-1})^{\nu-1} \\ &= L_i\,(1-p_i)^{\nu-1} - L_{i-1}(1-p_{i-1})^{\nu-1}, \end{aligned}$$

jossa $\Delta p_i = p_i - p_{i-1}$ ja toinen yhtäsuuruus on suora seuraus Lorenz-ordinaattojen määritelmästä.

Sijoitettaessa saadut tulokset laajennetun Gini-kertoimen lausekkeeseen

$$\begin{aligned} EG(\nu) &= 1-\nu\sum_{j=1}^{n} L_j\,(1-p_j)^{\nu-1}+\nu\sum_{j=1}^{n} L_{j-1}(1-p_{j-1})^{\nu-1} \\ &\quad -\sum_{j=1}^{n} r_i\,[(1-p_j)^{\nu}-(1-p_{j-1})^{\nu}]\,, \end{aligned}$$

havaitaan että ensimmäinen ja toinen summatermi kumoavat toisensa, sillä $L_0 = 0$ ja $p_n = 1$. Näin lopulliseksi tulokseksi saadaan haluttu muoto

$$EG(\nu) = 1+\sum_{j=1}^{n} r_j\,[(1-p_j)^{\nu}-(1-p_{j-1})^{\nu}].$$

Jakson 3.2.4 laskut

Merkitään suhteellisten osuuksien funktiota yleisesti h:ksi. Tällöin entropia-indeksi voidaan määrittää Theilin esittämällä tavalla

$$T^* = \sum_{j=1}^{n} f_j\frac{1}{N}h(\frac{1}{N})-\sum_{j=1}^{n} f_j\frac{r_j}{N}h(\frac{r_j}{N}).$$

eli on ilmeistä, että voidaan kirjoittaa

$$T^* = h(\frac{1}{N})-\sum_{j=1}^{n} w_j r_j h(\frac{r_j}{N}).$$

Theilin indeksissä $h = -\log$, joten saadaan

$$\begin{aligned} T &= \log(N)+\sum_{j=1}^{n} w_j r_j\log(r_j)-\log(N)\sum_{j=1}^{n} w_j r_j \\ &= \sum_{j=1}^{n} w_j r_j\log(r_j), \end{aligned}$$

sillä $\sum_j w_j r_j = 1$. Yleistetty entropiamitta saadaan, kun T^*:n yhtälöön sijoitetaan funktio

$$h(x;\alpha) = \frac{1-x^{\alpha-1}}{\alpha(\alpha-1)N^{1-\alpha}}.$$

Tällöin indeksiksi saadaan

$$\begin{aligned} GE(\alpha) &= \frac{1}{\alpha(\alpha-1)N^{1-\alpha}}\left[1-\frac{1}{N^{\alpha-1}}-\sum_{j=1}^{n} w_j r_j\left(1-(r_j/N)^{\alpha-1}\right)\right] \\ &= \frac{1}{\alpha(\alpha-1)N^{1-\alpha}}\left[1-\frac{1}{N^{\alpha-1}}-1+\frac{1}{N^{\alpha-1}}\sum_{j=1}^{n} w_j r_j r_j^{\alpha-1}\right] \\ &= \frac{1}{\alpha(\alpha-1)}\left[\sum_{j=1}^{n} w_j r_j^{\alpha}-1\right], \end{aligned}$$

kun $\alpha \neq 0$ ja $\alpha \neq 1$.

On ilmeistä, että parametrin arvot $\alpha = 0$ ja $\alpha = 1$ vaativat raja-arvotarkastelun. Yhtälöä tarkastelemalla on yhtä ilmeistä, että molemmat raja-arvot ovat 0/0-tyyppiä, jolloin voidaan turvautua L'Hôpital'n sääntöön eli tutkia osoittajan ja nimittäjän derivaattojen suhteen raja-arvoa. Derivaatta lasketaan rajankäyntimuuttujan suhteen.

Nimittäjän derivaatta on

$$2\alpha - 1.$$

Osoittajan derivaatan laskemiseksi osoittajan termit kannattaa kirjoittaa muodossa

$$r_j^\alpha = e^{\alpha \log(r_j)},$$

jolloin derivaatan laskeminen on suoraviivaista. Näin raja-arvo voidaan kirjoittaa muodossa

$$\lim_{\alpha \to 0 \text{ tai } 1} GE(\alpha) = \lim_{\alpha \to 0 \text{ tai } 1} \frac{1}{2\alpha - 1} \sum_{j=1}^{m} w_j \log(r_j) r_j^\alpha.$$

Tällöin on ilmeistä, että

$$GE(0) = -\sum_{j=1}^{n} w_j \log(r_j),$$

ja

$$GE(1) = \sum_{j=1}^{n} w_j \, r_j \log(r_j).$$

Osoitetaan vielä L'Hôpital'n säännön avulla, että edellä määritellyn funktion h raja-arvo parametrin arvolla $\alpha = 1$ on $-\log$ ja että $\lim_{x\to 0} x \log(x) = 0$.

Raja-arvo $\lim_{\alpha\to 1} h(x; \alpha)$ on ilmeisesti muotoa 0/0 eli L'Hôpital'n sääntöä voidaan käyttää. Nimittäjän derivaatta on muotoa

$$(\alpha - 1)c(\alpha) + \alpha N^{1-\alpha},$$

jossa apufunktio c saa äärellisen arvon, kun $\alpha = 1$. Osoittajan derivaatta puolestaan on aiemmin lasketun perusteella

$$-x^{\alpha-1} \log(x).$$

Raja-arvoksi saadaan siis

$$\lim_{\alpha \to 1} h(x; \alpha) = \frac{-x^0 \log(x)}{0 \cdot c(1) + 1} = -\log(x).$$

Raja-arvon $\lim_{x\to 0} x \log(x)$ laskeminen on suoraviivaista

$$\lim_{x \to 0} x \log(x) = \lim_{x \to 0} \frac{\log(x)}{1/x},$$

joka on muotoa $-\infty/\infty$. L'Hôpital'n sääntöä voidaan käyttää myös tällaisissa tilanteissa, eli

$$\lim_{x \to 0} x \log(x) = \lim_{x \to 0} \frac{1/x}{-1/x^2} = \lim_{x \to 0} -x = 0.$$

Liite B: Ei-positiivisten havaintojen suhteelliset osuudet

Taulukko B.1 Ei-positiivisten havaintojen suhteelliset osuudet (%) tarkastelun kohteena olevissa kulutustutkimuksissa.

Tulotyyppi	tulon määrä	1985	1990	1995	1998	2001
palkkatulot	nolla	13	19	24	25	24
	neg.	0,4	$< 0{,}1$	0,1	$< 0{,}1$	0,2
yrittäjätulot	nolla	16	37	47	28	26
	neg.	1,1	2,3	2,4	2,2	2,9
ansiotulot	nolla	3,4	8,0	12	8,2	7,3
	neg.	0,3	0,7	1,1	0,8	1,3
omaisuustulot	nolla	11	10	25	25	23
	neg.	15	15	8,4	5,0	3,0
tuot. tekijätulot	nolla	1,0	2,1	4,9	3,6	3,1
	neg.	0,4	0,7	0,6	0,4	0,4
saadut siirrot	nolla	9,3	13	8,8	12	14
	neg.	0	0	0	0	0,5
bruttotulot	nolla	0	0	$< 0{,}1$	0	$< 0{,}1$
	neg.	0	$< 0{,}1$	0	$< 0{,}1$	0
maksetut siirrot	nolla	3,9	3,8	2,3	2,3	2,0
	neg.	0	0	0	0	0
käyt. olevat tulot	nolla	0	0	$< 0{,}1$	0	$< 0{,}1$
	neg.	$< 0{,}1$	$< 0{,}1$	$< 0{,}1$	$< 0{,}1$	$< 0{,}1$
veropuhdistettu tulo	nolla	0	0	0	0	0
	neg.	0,3	0,4	0,3	0,3	0,4
havaintojen lkm		8200	8258	6743	4359	5495

Liite C: Tilastollisen analyysin taulukot

Taulukko C.1 Laajennetun Gini-kertoimen $EG(y;\nu)$ arvot sekä bootstrap-keskivirheet (suluissa). Bootstrap-otosten lukumäärä on 200.

	ν	1985	1990	1995	1998	2001
Tuotannontekijätulot	1,25	14,2	15,6	18,8	19,1	19,5
		(0,2)	(0,3)	(0,4)	(0,9)	(0,6)
	2	38,7	41,4	48,7	48,0	47,8
		(0,4)	(0,5)	(0,5)	(1,0)	(0,7)
	4	66,1	68,9	77,1	75,7	74,3
		(0,5)	(0,5)	(0,5)	(0,7)	(0,6)
	10	88,6	90,0	94,2	93,5	92,2
		(0,4)	(0,4)	(0,3)	(0,4)	(0,3)
Bruttotulot	1,25	9,8	10,3	11,0	12,2	13,4
		(0,2)	(0,3)	(0,4)	(0,7)	(0,5)
	2	25,4	26,0	27,2	29,3	31,5
		(0,3)	(0,4)	(0,5)	(0,8)	(0,7)
	4	41,5	41,7	42,6	45,5	47,8
		(0,3)	(0,4)	(0,5)	(0,7)	(0,6)
	10	55,5	55,1	55,5	58,9	61,2
		(0,4)	(0,4)	(0,5)	(0,7)	(0,6)
Käytettävissä olevat tulot	1,25	7,7	8,1	8,9	10,3	11,4
		(0,1)	(0,2)	(0,4)	(0,8)	(0,5)
	2	20,4	20,9	22,0	24,7	27,0
		(0,2)	(0,3)	(0,5)	(0,9)	(0,7)
	4	34,3	34,3	35,0	38,7	41,8
		(0,3)	(0,4)	(0,5)	(0,9)	(0,7)
	10	48,1	46,9	47,2	51,2	55,0
		(0,5)	(0,5)	(0,6)	(0,9)	(0,7)

Taulukko C.2 Testisuureen $Z_{t_1 \to t_2}$ arvot sekä tämän kaksisuuntaisen testin p-arvot (standardinormaalijakauma-approksimaatio).

	ν	$85 \to 90$	$90 \to 95$	$95 \to 98$	$98 \to 01$
Tuotannontekijätulot	1,25	3,7	6,0	0,3	0,4
		0,000	0,000	0,757	0,685
	2	4,3	9,9	−0,6	−0,1
		0,000	0,000	0,532	0,899
	4	3,8	11,3	−1,7	−1,6
		0,000	0,000	0,097	0,109
	10	2,6	9,3	−1,6	−2,5
		0,010	0,000	0,112	0,011
Bruttotulot	1,25	1,6	1,6	1,5	1,5
		0,112	0,102	0,135	0,140
	2	1,3	1,9	2,2	2,1
		0,177	0,055	0,025	0,040
	4	0,5	1,4	3,3	2,4
		0,616	0,153	0,001	0,015
	10	−0,7	0,5	3,9	2,6
		0,500	0,623	0,000	0,009
Käytettävissä olevat tulot	1,25	1,9	1,7	1,6	1,3
		0,062	0,082	0,118	0,206
	2	1,4	1,8	2,5	2,0
		0,154	0,076	0,011	0,043
	4	0,0	1,1	3,6	2,8
		0,998	0,252	0,000	0,006
	10	−1,7	0,4	3,9	3,3
		0,092	0,720	0,000	0,001

Liite D: Välillisten verojen luokat ja veroprosentit

Taulukko D.1 Välilliset veroprosentit COICOP-hyödykeluokissa.

Koodi	Selite	1985	1990	1995	1998	2001
C011	Elintarvikkeet	17,91	17,95	14,70	15,07	15,07
C012	Alkoholittomat juomat	20,11	20,14	17,08	17,61	17,61
C021	Alkoholijuomat	75,61	75,63	70,12	64,51	64,51
C022	Tupakka	71,30	71,34	66,33	55,81	55,81
C031	Vaatetus	16,84	17,05	18,06	17,96	17,96
C032	Jalkineet	18,11	18,49	19,52	18,98	18,98
C043	Asunnon huolto ja korjaus	16,08	16,12	17,14	18,08	18,08
C044	Muut asumispalvelut	11,16	11,16	11,88	14,36	14,36
C045	Sähkö, kaasu ja muut polttoaineet	35,63	35,63	34,40	39,58	39,58
C051	Huonekalut, matot ja sisustus	14,09	14,11	15,00	16,31	16,31
C052	Kodintekstiilit	17,72	17,98	19,03	18,47	18,47
C053	Kodinkoneet	17,04	17,09	18,16	18,07	18,07
C054	Lasiesineet, astiat ja taloustavarat	17,00	17,04	18,12	18,14	18,14
C055	Kodin työkalut ja laitteet	17,06	17,11	18,19	18,11	18,11
C056	Kodinhoitotavarat ja palvelut	14,38	14,39	15,31	16,91	16,91
C061	Lääkintätuotteet	21,10	21,12	20,99	21,70	21,70
C071	Ajoneuvojen hankinta	38,54	38,68	37,37	41,81	41,81
C072	Yksityisajoneuvojen käyttö	43,55	43,56	41,56	42,39	42,39
C073	Liikennepalvelut	8,46	8,46	4,90	7,04	7,04
C081	Tietoliikenne	16,91	16,92	21,74	18,01	18,01
C091	AV laitteet ja tietokoneet	17,28	17,41	18,47	18,17	18,17
C092	Muut suuret vapaa-ajan väl.	17,06	17,19	18,24	18,14	18,14
C093	Muut harrastusvälineet	16,34	16,41	17,43	17,07	17,07
C094	Virkistys ja kulttuuri	39,58	39,58	36,40	42,55	42,55
C095	Sanomalehdet ja kirjat	7,95	7,96	8,47	11,28	11,28
C096	Valmismatkat	6,11	6,11	6,50	10,66	10,66
C100	Koulutus	3,09	3,09	3,29	8,12	8,12
C111	Ravitsemispalvelut	16,60	16,60	17,66	18,00	18,00
C112	Majoituspalvelut	0,00	0,00	6,36	11,76	11,76
C121	Henkilökohtainen hygienia	16,90	16,91	17,99	18,02	18,02
C123	Henkilökohtaiset tavarat	17,59	17,81	18,87	18,58	18,58
C125	Vakuutukset	31,67	31,67	28,12	34,85	34,85
C126	Rahoituspalvelut	10,29	10,29	9,83	20,62	20,62
C127	Muut palvelut	8,68	8,68	9,24	12,64	12,64

Lähde: Jäntti (2004).

VATT-TUTKIMUKSIA -SARJASSA ILMESTYNEITÄ

PUBLISHED VATT RESEARCH REPORTS

91. Berghäll Elina – Heikkilä Tuomo – Hjerppe Reino – Kiander Jaakko – Kilponen Juha – Lavrac Vladimir – Stanovnik Peter: The Role of Science and Technology Policy in Small Economies. Helsinki 2002.
92. Räisänen Heikki (toim.): Rakenteellinen työttömyys. Tutkimusinventaari ja politiikkajohtopäätökset. Helsinki 2002.
93. Moisio Antti: Essays on Finnish Municipal Finance and Intergovernmental Grants. Helsinki 2002.
94. Parkkinen Pekka: Hoivapalvelut ja eläkemenot vuoteen 2050. Helsinki 2002.
95. Junka Teuvo: Maailman kilpailukykyisin maa? Tuottavuus ja investoinnit Suomessa 1975-2000. Helsinki 2003.
96. Cogan Josehp – McDevitt James: Science, Technology and Innovation Policies in Selected small European Countries. Helsinki 2003.
97. Perrels Adriaan – Kemppi Heikki: Liberalised Electricity Markets – Strengths and Weaknesses in Finland and Nordpool. Helsinki 2003.
98. Sarvimäki Matti: Euroopan Unionin itälaajentuminen ja maahanmuutto Suomeen. Helsinki 2003.
99. Räty Tarmo – Luoma Kalevi – Mäkinen Erkki – Vaarama Marja: The Factors Affecting the Use of Elderly Care and the Need for Resources by 2030 in Finland. Helsinki 2003.
100. van Beers Cees: The Role of Foreign Direct Investments on Small Countries' Competitive and Technological Position. Helsinki 2003.
101. Kangasharju Aki: Maksaako asumistuen saaja muita korkeampaa vuokraa? Helsinki 2003.
102. Honkatukia Juha – Forsström Juha – Tamminen Eero: Energiaverotuksen asema EU:n laajuisen päästökaupan yhteydessä. Loppuraportti. Helsinki 2003.
103. Simai Mihály (ed.): Practical Guide for Active National Policy Makers – what Science and Technology Policy Can and Cannot Do? Helsinki 2003.
104. Luoma Arto – Luoto Jani – Siivonen Erkki: Growth, Institutions and Productivity: An empirical analysis using the Bayesian approach. Helsinki 2003.
105. Montén Seppo – Tuomala Juha: Muuttoliike, työssäkäynti ja työvoimavarat Uudellamaalla. Helsinki 2003.
106. Venetoklis Takis: An Evaluation of Wage Subsidy Programs to SMEs Utilising Propensity Score Matching. Helsinki 2004.
107. Räisänen Heikki: Työvoiman hankinta julkisessa työnvälityksessä. Helsinki 2004.
108. Romppanen Antti: Vakaus- ja kasvusopimuksen ensimmäiset vuodet. Helsinki 2004.
109. Vaittinen Risto: Trade Policies and Integration – Evaluations with CGE Models. Helsinki 2004.

110. Hjerppe Reino – Kiander Jaakko (eds.): Technology Policy and Knowledge-Based Growth in small Countries. Helsinki 2004.

111. Sinko Pekka: Essays on Labour Taxation and Unemployment Insurance. Helsinki 2004.

112. Kiander Jaakko – Martikainen Minna – Voipio Iikko: Yrittäjyyden tila 2002-2004. Helsinki 2004.

113. Kilponen Juha – Santavirta Torsten: Competition and Innovation – Microeconomic Evidence Using Finnish Data. Helsinki 2004.

114. Kiander Jaakko – Venetoklis Takis: Spending Preferences of Public Sector Officials. Survey Evidence from the Finnish Central Government. Helsinki 2004.

115. Hämäläinen Kari – Ollikainen Virve: Differential Effects of Active Labour Market Programmes in the Early Stages of Young People's Unemployment. Helsinki 2004.

116. Räisänen Heikki: Recent Labour Market Developments in Europe. Helsinki 2005.

117. Ropponen Olli: Kokonaiskulutuksen kehitys Suomessa talouden ulkopuolisten tekijöiden suhteen vuosina 1985–2001. Helsinki 2005.

118. Räty Tarmo – Luoma Kalevi – Aaltonen Juho – Järviö Maija-Liisa: Productivity and Its Drivers in Finnish Primary Care 1988–2003. Helsinki 2005.

119. Kangasharju Aki – Aaltonen Juho: Kunnallisen päivähoidon yksikkökustannukset: Miksi kunnat ovat niin erilaisia? Helsinki 2006.

120. Perrels Adriaan – Ahlqvist Kirsti – Heiskanen Eva – Lahti Pekka: Kestävän kulutuksen mahdollisuudet ekotehokkaassa elinympäristössä. Helsinki 2006.

121. Berghäll Elina – Junka Teuvo – Kiander Jaakko: T&K, tuottavuus ja taloudellinen kasvu. Helsinki 2006.

122. Rauhanen Timo – Peltoniemi Ari: Elintarvikkeiden ja ruokapalveluiden arvonlisäverotus EU:ssa ja Suomessa. Helsinki 2006.

123. Kiander Jaakko – Martikainen Minna – Pihkala Timo – Voipio Iikko: Yritysten toimintaympäristö: Kyselytutkimuksen tuloksia vuosilta 2002–2005. Helsinki 2006.

124. Räty Tarmo – Kivistö Jussi: Mitattavissa oleva tuottavuus Suomen yliopistoissa. Helsinki 2006.

125. Teppala Tiina: Kulutusverotus teoriasta käytäntöön – Vaikuttaako arvonlisäverotus kuluttajahintoihin? Helsinki 2006.

126. Ulvinen Hanna: Suomen elintarvike- ja ruokapalvelualan rakenne, kilpailullisuus ja taloudellinen suorituskyky. Helsinki 2006.

127. Aaltonen Juho – Kirjavainen Tanja – Moisio Antti: Efficiency and Productivity in Finnish Comprehensive Schooling 1998–2004. Helsinki 2006.

128. Mattila-Wiro Päivi: Changes in the Distribution of Economic Wellbeing in Finland. Helsinki 2006.

129. Kiander Jaakko: Julkisen talouden liikkumavara vuoteen 2030 mennessä. Helsinki 2007.