

TAMPEREEN YLIOPISTO

Taloustieteiden laitos

**ALTRUISMI, ITSEKKYYS JA
RATIONAALINEN MAKSIMOINTI**

Kansantaloustiede

Pro gradu –tutkielma

Toukokuu 2010

Ohjaajat: Jari Vainiomäki

ja Kaisa Kotakorpi

Mika Niemi

TIIVISTELMÄ

Tampereen yliopisto:	Taloustieteiden laitos, Kansantaloustiede
Tekijä:	NIEMI, MIKA
Tutkielman nimi:	Altruismi, itsekkyyden ja rationaalinen maksimointi
Pro gradu -tutkielma:	72 sivua
Aika:	Toukokuu 2010
Avainsanat:	Altruismi, Homo Economicus, egoismi, rationaalisuus, tieteenteoria, hyötyfunktio, yhteishyväpeli

Taloustieteen perusparadigma nojaa ihmisen rationaaliseen itsekeskeisyyteen – *Homo Economicukseen* – taloudellisen tehokkuuden selittäjänä. Tämä puhtaasti taloudellisten ja itsekkäiden hyötyjen tavoittelu ihmisluonnon perustana on aina herättänyt vahvaa vastarintaa etenkin sosiologian ja filosofian saroilla. Niin sosiologiassa, antropologiassa kuin biologiassakin korostetaan myös muita inhimillistä toimintaa ohjaavia insentiivejä kuten altruismia. Myös taloustieteen behavioristiset tutkimukset ovat löytäneet paljon näyttöä altruistisista insentiiveistä. Onko altruismia siis olemassa, kuinka se tulisi taloustieteissä määritellä ja onko se ristiriidassa taloustieteen peruspilarin Homo Economicuksen kanssa?

Tavoitteeni on tarkastella altruististen insentiivien luonnetta, todisteita ja suhdetta toiminnan rationaalisuuteen. Pyrin osoittamaan, että altruistista käyttäytymistä todella on olemassa, altruistinen käyttäytyminen voidaan nähdä rationaalisena toimintana ja altruismi on mahdollista huomioida käytännön tasolla taloustieteen mallinnuksissa, sovelluksissa ja hyötyfunktioissa.

Oletukseni on, että esittämäni ongelman ratkaisun avaimet piilevät lähinnä taloustieteen tieteenteorian täsmällisessä ja selkeässä määrittelyssä. Tulen sekä tarkastelemaan perinteistä semantiikkaa että esittämään vaihtoehtoisen tavan jäsentää taloustieteen perusparadigma huomioiden myös altruistiset insentiivit.

SISÄLLYSLUETTELO

1.	Johdanto.....	1
2.	Altruismin ja Homo Economicuksen vastakkainasettelu	5
2.1.	Homo Economicus –oletuksen ja altruismin ristiriita	5
2.2.	Altruismi	9
2.3.	Rationaalisuus	11
2.4.	Rationaalisuusoletuksen tieteenteoreettista taustaa	13
3.	Kokeet altruismin empiirisinä todisteina.....	15
3.1.	Perinteiset pelien elementit	15
3.1.1.	Uhkavaatimuspeli.....	16
3.1.2.	Ulkopuolinen rankaisija –peli	17
3.1.3.	Diktaattoripeli	18
3.1.4.	Kritiikkiä peleistä	19
3.2.	Yhteishyväpeli.....	20
3.2.1.	Yhteishyväpelin rakenne	20
3.2.2.	Vahva vastavuoroisuus ja talouden toimijan stereotyypit	25
3.2.3.	Altruistiset ja egoistiset stereotyypit yhteishyväpelissä	27
3.2.4.	Altruismin yhteiskunnallinen merkitys – evolutiivinen malli	29
3.3.	Sosiaalinen vaihdanta ja vahva vastavuoroisuus.....	37
3.3.1.	Sosiaalisen vaihdannan kokeet – työmarkkinat	38
3.3.2.	Sosiaalisen vaihdannan seuraukset	39
3.4.	Altruismin neurotaloustieteellinen perusta	42
4.	Onko altruismi rationaalista?.....	45
4.1.	Altruistisen hyötyfunktion mahdollisuus	45
4.2.	Adam Smithin ongelma	54
5.	Teorioita altruistisista hyötyfunktioista.....	58
5.1.	Sosiaalisten preferenssien malli	58
5.2.	Toisistaan riippuvaiset preferenssit	63
5.3.	Aikeisiin perustuva vastavuoroisuus	63
6.	Johtopäätöksiä.....	65
	LÄHTEET	68
	LIITE 1.	72

1. Johdanto

Tutkielmassaan *The Wealth of Nations* (suom. *Kansojen varallisuus*) (1776) Adam Smith painottaa yksilön omanvoitonhakuisuuden tärkeyttä, mikä johtaa loppupäätelmään kilpailullisten markkinoiden tehokkuudesta. Taloustieteen perusparadigma nojaa vielä nyt kahdensadan vuoden jälkeenkin ihmisen rationaaliseen itsekeskeisyyteen – *Homo Economicukseen* – taloudellisen tehokkuuden selittäjänä (Solarz, 2006). Homo Economicus viittaa siis yksilön rationaaliseen toimintaan, joka klassisessa taloustieteessä on vahvasti liitetty toimintaan taloudellisten, itsekkäiden, yksityisten ja instrumentalististen tarkoitusten saavuttamiseksi (Zafirovski, 2003). Kuitenkin monissa muissa tieteissä aina sosiologiasta biologiaan korostetaan myös muita inhimillistä toimintaa ohjaavia insentivejä kuten altruismia. Altruistinen toiminta on kaikkea muuta kuin taloudellisten tai itsekkäiden tarkoitusprien tavoittelua – se on yhteisöllistä ja epäitsestä toimintaa. Onko altruismia siis olemassa, kuinka se tulisi taloustieteissä määritellä ja onko se ristiriidassa taloustieteen peruspilarin Homo Economicuksen kanssa?

Altruismi on ollut ihmisen toiminnan taloustieteellisen selittämisen kompastuskiviä, jonka ratkaisemiseksi on ehdotettu kääntymään ihmistieteiden puoleen. Perinteisesti sen käsittely taloustieteissä on täysin sivuutettu tai sitä on vältelty. Tämä vaikuttaa johtuneen lähinnä epämääräisestä semantiikasta, joka on muodostunut Homo Economicus -oletuksen, altruismin ja rationaalisuuden sekä näiden keskinäisten suhteiden ympärille. Rationaalisuuden käsite tuntuu kirjallisuudessa usein

sekoittuneen egoismin käsitteen kanssa, ja siten näkemys ristiriitaisuudesta altruismin kanssa on vahvistunut. Tähän sekaannukseen ja sen syihin palataan kappaleessa 2.3.

Tämänhetkiset geenievoluutioon perustuvat teoriat eivät kykene selittämään altruistisen käyttäytymisen piirteitä, mikä viittaa kulttuurievoluutioteorioiden tärkeyteen ilmiön selittämisessä (Fehr ja Fischbacher, 2003; Fehr ja Gächter, 2002). Tämä tarkoittaa, että yksilön toimintaa ei voida nähdä yksilössä absoluuttisesti geneettisten ominaisuuksien summana vaan selittämisessä pitää huomioida myös yhteisöllisyyden sosiologiset vaikutukset. Pitäisikö siis myös taloustieteellisessä selittämisessä siirtyä relativistisempiin, inhimilliset tunteet ja motivaatiot huomioiviin lähtökohtiin (esim. behavioristiset teoriat), kuten evolutiivisessa tutkimuksessa on tehty?

Tämä taloustieteen kylmän rationaalinen, instrumentalisesti hyvinvointiaan maksimoiva Homo Economicus tieteenteorian keskiössä on myös herättänyt jo pitkään vahvaa vastustusta muissa tieteenhaaroissa. Jopa taloustieteen sisällä on ollut behavioristista liikettä, joka on paitsi voimakkaasti kritisoinut perusaksioomaa puhtaasti itsekkäästä toimijasta myös tarjonnut paljon todisteita epäitsekkäiden intentioiden puolesta. Tästä huolimatta valtavirtataloustiede on edelleen rakentunut yksipuolisen ja virheellisen ihmiskuvan varaan (Fehr ja Gintis, 2007).

Tästä ei kuitenkaan voida yksin syyttää taloustiedettä, jonka tutkimuskohteena pääasiassa on talouden toiminta yhteiskunnan näkökulmasta. Tällöin myös mikrotason tarkastelut painottuvat lähinnä yksilöiden agregattiin – eivät yksilön intentioihin. Sen sijaan tieteenvälisen tehtävänjakojen mukaan yhteisössä toimivien yksilöiden ja heitä yhteisöllisessä toiminnassa ajavien intentioiden tutkimuksen sekä

määrittelyyn on ajateltu perinteisesti langenneen sosiologian vastuulle. Sosiologia on kuitenkin epäonnistunut tehtävässään tarjota taloustieteelle hyvä lähtökohta yksilön taloudellisen toiminnan tarkasteluun. Se ei ole onnistunut luomaan yhtä yleisesti hyväksyttyä teoriaa, joka selittäisi yhteiskunnassa näkyvien toimijoiden aggregaatiot yksilöiden preferenssien, uskomusten, sosiaalisten ja taloudellisten rajoitteiden sekä kannustimien pohjalta (Fehr ja Gintis, 2007). Koska sosiologia ei kykene tukemaan taloustieteitä tieteenteoreettisella tasolla, tieteenhaaran on kannettava oma vastuunsa aksiomistaan. Näin ollen myös altruismin tutkimus on tullut osaksi taloustieteen pohdintoja.

Tavoitteeni on tarkastella altruististen insentiivien luonnetta, todisteita ja suhdetta toiminnan rationaalisuuteen. Näkemykseni mukaan perusteet kritiikille rationaalista ja itsestä Homo Economicus –oletusta vastaan löytyvät, mikäli:

1. Altruistista käyttäytymistä todella on olemassa,
2. Altruistinen käyttäytyminen voidaan nähdä rationaalisena toimintana ja
3. Altruismi on mahdollista huomioida käytännön tasolla taloustieteen mallinnuksissa, sovelluksissa ja hyötyfunktioissa.

Väitteeni on, että kaikki yllä nimeämäni ehdot voidaan osoittaa toteen. Toisaalta väitän, että näin oikeutettu kritiikki voidaan ohittaa suhteellisen hyvin, mikäli Homo Economicus -oletus määritellään uudelleen aiempaa laajemmaksi koskemaan myös epäinstrumentalististen hyötyjen maksimointia. Oletukseni siis on, että esittämäni ongelman ratkaisun avaimet piilevät lähinnä taloustieteen tieteenteorian täsmällisessä ja selkeässä määrittelyssä.

Tavoitteena on tutkia altruismia niin taloustieteen filosofian, käytännön ja empiiristen kokeiden, taloustieteen sovellusten kuin teoriansakin näkökulmista. Ensin kappaleessa kaksi selvitän ja jäsenän tarkemmin esittämäni ongelman tieteenteoreettista taustaa. Tämä tarkoittaa myös osaltaan altruismin ja rationaalisuuden käsitteiden täsmällistä määrittelyä. Kappaleessa kolme esitän empiirisiä todisteita altruististen insentiivien olemassaolosta niin taloustieteen kuin muidenkin tieteiden näkökulmasta. Kun altruismi on kyetty näkemään osana luonnollista yksilön käytöstä ja sen luonne on määritelty tarkemmin, tutkitaan tämän toiminnan suhdetta rationaaliseen toimintaan. Kappaleessa neljä pyritään osoittamaan altruismin olevan nähtävissä myös rationaaliseksi koettavana toimintana. Kun altruismi voidaan nähdä rationaalisena toimintana, sen pohjalta on lopulta mahdollista rakentaa myös hyötyfunktioita. Kappaleessa viisi esitellään näistä muutamia. Työn lopussa on Liite 1, joka sisältää määrittelyjä keskeisistä filosofisista termeistä, sillä työn kysymyksenasettelu linkittyy vahvasti tieteenteoriaan sekä filosofiaan. Näin ollen asioiden käsittely myös filosofisin termein on välttämätöntä.

2. Altruismin ja Homo Economicuksen vastakkainasettelu

Tässä kappaleessa esittelen ensin tarkemmin tarkasteltavan ongelman käsitteellisellä tasolla. Tämän jälkeen määrittelen altruismin ja rationaalisuuden termit siten kuin tulemme ne tässä tarkastelussa ymmärtämään. Lisäksi käsittelen lopuksi rationaalisuusoletusta sillä uskon sen olevan taloustieteissä usein epämääräisesti määritelty.

2.1. Homo Economicus -oletuksen ja altruismin ristiriita

Kysymyksenasettelu ei suinkaan ole yksinkertainen. Ensinnäkin teoriat ihmisen toiminnan perimmäisistä intentioista jakaantuvat tieteenalottain hyvin laajalle aina sosiologiasta ja mielen filosofiasta teologiaan. Tieteellistä konsensusta ei tähän liene muutoinkaan edes mahdollista saavuttaa. Vaikka myöhemmin käsiteltävät neurologian tulokset antavatkin tähän joitain tieteellisesti mitattuja vihjeitä, perimmäiset intentiot jäävät aina uskojen ja oletusten varaan. Toiseksi taloustiede on aiheeseen liittyvässä kirjallisuudessa pitkälti laiminlyönyt tieteenteoreettisten lähtökohtien nimeämisen ja siihen liittyvät semanttiset valinnat. Tästä johtuen altruismin, rationalisuuden ja egoismin termejä on käytetty räikeästi eri merkityksissä ja toisinaan jopa sekoitettu toisiinsa. Homo Economicus puolestaan on kasvanut ilmiöksi, jota ei ole nähty tarpeelliseksi selittää tai avata, vaikka se samalla on merkitykseltään jatkuvasti jalostunut ja kehittynyt.

Näistä syistä työni ensimmäiseksi haasteeksi nousee määrittellä altruismin, rationaalisuuden, egoismin ja Homo Economicuksen termit. Tätä tehdessä kärjistän taloustieteen näkemyksen perinteisen talusteorian mukaiseksi, vaikka vaihtoehtoisiaakin näkemyksiä toki on esitetty. Vastaavasti pyrin halki työn käyttämään

samoja valittuja termejä viitatessani perinteisiin määritelmiin. Lisäksi tulen tuomaan rinnalle uusia termejä aina kun esittelen uuden näkemyksen em. ilmiöiden olemuksesta. Väistämättä tarkoituksena kuitenkin on määritellä perinteisiä termejä uudestaan, joten työn lopussa semantiikka tarkoitushakuisesti syventyy ja laajentuu alkuperäisten avainsanojen osalta.

Varsinainen ongelma altruismin ja taloustieteen suhteesta periytyy jo sen oppi-isän Adam Smithin jakaantuneista näkemyksistä. Taloustieteen perustan rakentanut *The Wealth of Nations* maalaa kuvan Homo Economicuksesta, joka puolestaan on omaksuttu taloustieteellisen ihmiskuvan pohjaksi. Toisaalta myös jo Smith näki toisenkin – altruistisen – puolen ihmisluonnosta. Tätä hän kuvasi toisessa pääteoksessaan *Theory of Moral Sentiments*. Näennäinen ristiriitaisuus tieteenteoriassa kuitenkin heti alusta lähtien ohitettiin ja hyväksyttiin yksipuolinen kuva Homo Economicuksesta. Tähän näennäiseen ristiriitaisuuteen palaan uudestaan luvussa 4.2.

On toki ymmärrettävää, etteivät ensimmäiset taloustieteilijät olleet niinkään kiinnostuneita perusparadigmansa määrittelystä kuin konkreettisista sovelluksista. Sen sijaan tuntuu yllättävältä, ettei tätä ongelmaa ole myöhemmin tuotu selkeämpänä kysymyksenä esille. Aiheen sivuuttaminen on johtanut siihen, että Homo Economicus on taloustieteen kielessä samaistettu puhtaasti rahallista hyötyään rationaalisesti maksimoivaan egoistiin. Epätäsmällisen terminologian ja ajoittain virheellisen arkikielen kautta rationaalisesta toiminnasta ja egoismista on myös tullut synonyymejä toisilleen. Rationaalisuus, egoismi ja Homo Economicus ovat tulleet yleiskäsitteiksi, joilla viitataan taloustieteen ihmiskuvaan. Koska altruismi on

käsitteellisesti egoismin vastakohta, se on edellämainitun virheellisen johtopäätelmän kautta mielletty myös rationaalisuuden vastakohdaksi. Tämä päättelyketju ei siten ole jättänyt tilaa altruismille taloustieteellisessä ihmiskuvassa.

Altruismin ongelma on kuitenkin ohittamaton. Yksilön toimintaa ei voida yleismaailmallisesti ja kaikissa tilanteissa kuvata ilman altruististen aspektien myöntämistä. Perinteinen taloustiede on kuitenkin pyrkinyt välttämään tämän ongelman väittäen myös altruistiseksi näyttäytyvien insentiivien pohjautuvan aina oman edun välilliseen ajamiseen.

Johdannossa jaoin tavoitteeni kolmeen osaan. Pyrin seuraavaksi avaamaan jokaisen kohdan perusteita.

1. Altruistista käyttäytymistä todella on olemassa.

Behavioristinen taloustiede on viime aikoina tehnyt entistä ahkerammin ja monipuolisemmin kokeita, jotka mallintavat yksilön taloudellista toimintaa ja intentioita. Näitä kokeita ja niiden sovelluksia käydään läpi kappaleessa 3. Tämä todistusaineisto antaa tukea altruismin olemassaololle. Samalla tullaan toteamaan altruismin olevan paitsi todellista myös *välttämätön* osa yhteiskunnan muodostumista ja toimintaa (Gintis, 2000, kappale 3.2.4). Tämä tarkastelu johtaa uuden kysymyksen äärelle. Kun altruismin olemassaolo näin myönnetään, joudutaan väistämättä ottamaan kantaa Homo Economicus -oletukseen.

- I. Homo Economicus -oletus on hylättävä taloustieteen paradigmana sen muodostaman virheellisen ihmiskuvan takia tai
- II. Oletusta on laajennettava ja sen termeissä käytettyä semantiikkaa tarkennettava.

Vaihtoehto II on näistä ensisijainen, sillä koko taloustieteen perinteen ja sen tulosten hylkääminen tai kyseenalaistaminen olisi luonnollisesti hyvin ongelmallista. Toisaalta se tuo uuden haasteen. Tieteenteoreettista pohjaa voi toki aina laajentaa, mutta sen vieminen liian yleismaailmalliseksi kadottaa sen merkityksen. Taloustieteen tavoitteena on tehdä konkreettisia ja eksakteja teorioita sekä malleja. Tästä syystä tarvitaan lisää ehtoja, jotta voidaan oikeuttaa Homo Economicus -oletuksen uudelleenmäärittely aiempaa laajemmalla merkityksellä.

2. Altruistinen käyttäytyminen voidaan nähdä rationaalisena toimintana.

Altruismin ja rationaalisuuden on kyettävä olemaan sopuinnussa toistensa kanssa, mikäli taloustieteen kaikkia aiempia saavutuksia ei olla valmiita kyseenalaistamaan. Hyvin laaja osa taloustieteen teorioista pohjaa rationaaliselle toiminnalle. Perinteinen talousteoria epähuomiossaan väittää, että altruismin hyväksyminen tarkoittaa rationaalisuuden hylkäämistä. Tällöin kuitenkin ei voitaisi pitää tarkoituksenmukaisena Homo Economicus -oletuksen muuttamista, sillä se muuttaisi liikaa koko tieteenalan hyviä ja laajan hyväksynnän saaneita perusrakenteita. Niinpä kappaleessa 4 tullaan osoittamaan altruismin olevan sopuinnussa rationaalisen käyttäytymisen kanssa.

3. *Altruismi on mahdollista huomioida käytännön tasolla taloustieteen mallinnuksissa, sovelluksissa ja hyötyfunktioissa.*

Kun kohtien 1 ja 2 jälkeen olemme voineet todeta Homo Economicus -oletuksen uudelleenmäärittelyn olevan paitsi tarpeellista myös mahdollista, se pitää todistaa myös hyödylliseksi. Altruismin hyväksymisen ja perusparadigman monimutkaistamisen täytyy kyetä tarjoamaan taloustieteelle mahdollisuus entistä parempiin teorioihin. Tämä tarkoittaa yksilön toiminnan parempaa selittämistä sekä mahdollisuutta kuvata sitä hyötyfunktioilla, joita voidaan soveltaa useille taloustieteen eri osa-alueille. Kappaleessa 5 esitellään siksi muutamia erilaisia teorioita altruistiset insentiivit huomioivista hyötyfunktioista.

Mikäli yllämainitut kohdat pystytyään tyydyttävästi osoittamaan toteen, on meillä väitteeni mukaan syy ehdottaa taloustieteen ihmisenäkemyksen ja perusparadigman uudelleen määrittelyä ja termien parempaa jäsentämistä.

2.2. Altruismi

Altruismin määritelmä on pitkälti periytynyt kantilaisesta filosofiasta. Tiukan kantilaisen määritelmän mukaan käyttäytyminen on ollut altruistista vain silloin, kun yksilö on kokenut siitä ainoastaan kärsimystä ja täten hyödyn – niin rahallisen kuin henkisenkin – laskua. Tähän toisaalta lienee liittynyt näkemys moraalien luonteesta. Kantin kategorisen imperatiivin mukainen velvollisuusetiikka ei näe moraalia valintana vaan velvollisuutena. (Kant, 1785 – 1788) Mikäli yksilö kokee oman moraalinsa seuraamisesta *sisäistä lämpöä* tai mielihyvää, ei teko ole ollut kantilaisittain altruistinen. Tämä määritelmä on kuitenkin hankala. Mikäli määritelmä pohjataan ihmismielen perimmäisiin intentioihin varsinaisten yksilön valintojen ja

toimien sijaan, se jää lähinnä filosofiseksi kielileikiksi. Taloustieteen termistössä mahdollisuus käytännön sovelluksiin on kuitenkin aina oltava.

Taloustieteen perinteinen altruismin määritelmä noudattelee pitkälti kantilaisia linjoja. Kuitenkin siinä, missä Kant ottaa esiin moraalin ja henkiset vaikutukset, taloustiede sivuuttaa jälleen kaikki epäinstrumentalistiset tekijät. Tämä ratkaisee edellämaintun ongelman kantilaisessa määritelmässä, mutta toisaalta se on melko vahva harppaus toiseen äärelaitaan.

Joka tapauksessa ymmärrän tässä työssä ”altruistiset motivaatiot” kuten altruismin tutkimuksessa taloustieteissä suuri joukko tutkijoista (esim. Elster, 2006). Altruistinen motivaatio on halua ja valmiutta kohottaa muiden hyvinvointia, vaikka se aiheuttaa yksilölle itselleen hyvinvoinnin nettolaskua. ”Altruistinen toiminta” puolestaan ymmärretään tässä toimintana, jolle altruistinen motivaatio tarjoaa riittävän syyn. Käytännössä tässä työssä – etenkin sen alussa – altruistiseksi kutsutaan toimintaa, joka sellaiseksi ulospäin näyttäytyy hyödyn ollessa instrumentalistista.

Edellinen määritelmä on kuitenkin täysin riippuvainen maksimoitavan hyödyn luonteesta. Jos hyötyyn, jota yksilö haluaa maksimoida, luetaan myös kokemus oman moraalin seuraamisen tärkeydestä, täytyy altruismi määritellä uudelleen. Mikäli lopulta nimittäin toteamme Homo Economicus -oletuksen uudelleenmäärittelyn olevan tarpeellista ja laajennamme hyödyn myös epäinstrumentalistisiin aspekteihin (mm. moraalit), ei yllä mainittu määritelmä välttämättä anna juurikaan tuloksia. Mikäli altruistiseksi tulkitsemamme tilanne antaa yksilölle henkisiä hyötyjä, olemme jälleen tilanteessa, jossa termit ovat vain filosofisen kielileikin kohteita eivätkä vastaa todellisuutta. Tässä tapauksessa altruismi tulee määritellä puhtaasti vain haluna

kohottaa muiden *instrumentalistista* hyvinvointia oman *instrumentalistisen* nettohyvinvoinninlaskun myötä.

Kaikki edellä mainittu siis liittyy edun määrittelyn ongelmaan taloudellisen hyödyn ja onnellisuuden välillä sekä näiden termien väliseen kuiluun, jonka ohittamisesta Solarz (2006) ekonomisteja syyttää. Vaikka edellä mainittu uudelleenmäärittely on näennäinen ratkaisu altruismin ja Homo Economicus –oletuksen ristiriidan välillä, on myös muistettava taloustieteen tärkeä rooli melko eksaktia tietoa tuottavana tieteenalana. Taloustieteen ei ole tarkoitus selittää kaikkea jotenkin, vaan jotakin (taloutta ja siihen liittyviä seikkoja) tarkasti. Tästä syystä on helppo ymmärtää, miksi – kuten Staveren (2001) esittää – modernistisen taloustieteen metodologia on rakennettu dualistisesti: hyöty määritellään taloudelliseksi, moraaliala ei; järkea pidetään osana taloustiedettä, tunteita ei. Tämän fundamentalistisen dualismin kritiikki kuitenkin tulee olemaan olennainen osa, kun lähdemme pohtimaan, kuinka taloudentoimijan käytöstä voidaan kuvata kokonaisvaltaisemmin altruistiset insenttiivit huomioiden.

2.3. Rationaalisuus

Rationaalisuus on määriteltävissä monella eri tavalla näkökannasta riippuen. Esimerkiksi itseään toteuttava ihminen voi mieltää toimintansa rationaaliseksi, kunhan hän toiminnallaan pääsee toteuttamaan itseään jopa ”mihin hintaan hyvänsä” (Pihlanto 2002, 75). Taloustieteen perinteinen määritelmä kuitenkin nojaa jo aiemmin mainittuun dualismiin taloudellisten termien sekä ei-taloudellisten termien välillä. Taloustieteiden näkökulmasta rationaalisuus liittyy siis puhtaasti materialistisen ja mitattavan hyvinvoinnin tavoitteluun. Tämä väitteeni pohjautuu huomioihin lähteinä

käytettyjen julkaisujen kielestä. Kuten jo aiemmin olen todennut, vaikuttaa siltä, että taloustiede on määrittelyt sivuuttaen rakentanut Homo Economicuksesta kokonaisvaltaisen ilmiön, johon myös rationaalisuus rinnastetaan.

Hyötyfunktioita tarkastellessa rationaalisuus on yleensä taloustieteissä helposti mielletävissä korkeimman hyötytason valintana. Ei ole rationaalista valita pistettä, joka jää valittavissa olevan budjettisuoran alapuolelle, vaan rationaalinen valinta on aina budjettisuoralla. Tätä tarkastelua syvennetään luvussa 4. Tässä valossa on kuitenkin hyvin nähtävissä tausta huomiolleni rationaalisuuden samaistuksesta vain instrumentalistisiin hyötyihin. Kun rationaalisuus on voitu liittää hyötyfunktio tarkasteluihin, se on kohdistunut vain instrumentalistisiin tekijöihin. On siis luonnollista, että siitä on muodostunut synonyymi egoismin kanssa. Näin myös altruismin ja rationaalisuuden välille on tullut ristiriita.

Rationaalisuuden määrittely suhteessa hyötyfunktio tarkasteluihin on kuitenkin hyvä. Mikäli joudumme määrittelemään altruismin, hyödyn ja Homo Economicuksen termejä uudelleen, kykenemme edelleen pitäytymään samassa rationaalisuuden määritelmässä. Jos epäinstrumentalistiset hyödyt voidaan mallintaa osaksi hyötyfunktiota, myös rationaalisuuden rajapinta on tällöin jo luonnollisesti laajentunut koskemaan niitä.

Taloustieteen valtavirtaa edustavassa uusklassisessa oppisuunnassa on myös valikoitunut vähittäisen kehityksen kuluessa eri vaihtoehdoista oletus rajoittumattomasta rationaalisuudesta. Siinä ihmisillä ajatellaan olevan täydellinen tai ainakin jäsentynyt tietämys kaikista päätöksiinsä vaikuttavista asioista. Jokainen tietää, mitä hän ei tiedä, ja pystyy arvioimaan tietämyksensä lisäämisen

kannattavuutta. Myös tiedon käsittelyn oletetaan olevan täydellistä. Tunteiden ei katsota vaikuttavan millään tavoin päätöksiin, tai ainakin niiden tarkastelua eksogeenisina pidetään riittävänä. (Vihanto 2005, 1.) Moraali ja sitä kautta mm. altruistinen insentiivi on kuitenkin pohjimmiltaan *tunne*, näin ollen rajoittumaton rationaalisuus niin ikään on ollut omiaan syventämään kuilua altruismin ja rationaalisuuden käsitteiden välillä.

2.4. Rationaalisuusoletuksen tieteenteoreettista taustaa

Pyrin seuraavassa kuvaamaan lyhyesti rationaalisuusoletuksen tieteenteoreettista taustaa omia ajatuksiani ja lähtökohtiani avatakseni. Tämän taustan yhteys arkipäivän taloustieteeseen ei ole suora ja siksi käsittelyyn tarvittava termistö on lainattu filosofiasta. Siitä huolimatta koen tämän taloustieteelle keskeisen rationaalisuusparadigman tieteenteoreettisen käsittelyn oleellisena tai peräti välttämättömänä taloustieteen perusluonteen ymmärtämiseksi. Rationaalisuusoletuksen taustaa voidaan valottaa tarkastelemalla sen takana olevia metodologisia valintoja, joiden voidaan nähdä pohjaavan luonnontiedelähtöiseen – positivistiseen – tiedekäsitykseen.

Objektivistinen lähestymistapa valitsee ontologiseksi oletukseksi realismin. Sen mukaan maailmalla on todellinen rakenteensa. Tämä voidaan paljastaa tieteellisen metodin – usein luonnontieteiden – avulla. Subjektivistinen lähestymistapa puolestaan otaksuu ihmisen luovan omia käsitteistöjään ja rakenteitaan pyrkiessään ymmärtämään maailmaa, joka konstruoituu hänen omasta tietoisuudestaan, kielestään ja kulttuuristaan. Tietoteoreettisen oletuksensa pohjaksi objektivismi valitsee positivismin ja näkee tieteellisen tiedonhankinnan kumulatiiviseksi toiminnan

selittämisen- ja ennakointiprosessiksi – paljolti luonnontieteiden tapaan. Subjektivistisen antipositivismin mukaan tieto taas on ihmisiin sitoutunutta ja sitä on näin ollen haettava heiltä itseltään. (Pihlanto 2002, 75.)

Valinta objektivismin ja subjektivismin välillä liittyy vahvasti siis käsitykseen ihmisestä. Jos maailma nähdään konkreettisenä ja mahdollisena objektiivisen havainnoinnin kohteena, se on ihmisestä riippumaton järjestelmä. Positivistin mielestä tällainen maailma on täysin selitettävissä objektiivisesti, kuten myös ihmisen toiminta, joka voidaan näin nähdä vaikkapa rationaaliseksi. Subjektivismissa maailma taas näyttäytyy sosiaalisena – ihmisten ja kulttuurin luomana konstruktiona. Subjektivismissa ihminen nähdään todellisuuden ja tiedon keskeisenä tuottajana, jonka määrittely on tällöin erityisen tärkeää. Kun ihmisen määrittely nousee keskeiseksi tekijäksi, on mietittävä, voidaanko todella käyttää Homo Economicus – oletusta, joka ei vastaa ihmisen todellista perusluontoa. Koska omat tieteenteoreettiset näkökantani pohjautuvat lähinnä fenomenologisille näkemyksille (fenomenologian mukaan todellisuus muotoutuu sen suhteesta ihmisen kokemusmaailmaan), joudun osittain hylkäämään Homo Economicus -oletuksen ja etsimään behavioristisille taloustieteen teorioille tyypillisempiä ratkaisuja. Lähtökohtaisesti en kuitenkaan kiellä altruistisen käyttäytymisen kumpuamista rationaalisesta egoismista, ja pyrin jopa tätä kantaa osittain edustamaan. Tämä kuitenkin vaatii laveampaa käsitystä niin ihmisen edusta kuin rationalismistakin.

3. Kokeet altruismin empiirisinä todisteina

Tässä kappaleessa käsittelen altruistista – perinteisesti irrationaaliseksi nähtyä – käytöstä empirian valossa esittämällä perusteita, jotka puhuvat irrationaalisen käytöksen sijaan altruististen preferenssien (other-regarding preferences) puolesta. Kappaleessa 3.1 esittelen muutamia yksinkertaisia koetuloksia, jotka viittaavat ihmisten välittävän myös toisten ihmisten hyvinvoinnista. Nämä tutkimustulokset antavat perustan etsiä altruistista hyötyfunktioita, joka voisi synnyttää havaitun käytöksen. Kappaleessa 3.2 käsittelen muita pelejä kompleksimpaa mutta myös enemmän sovelluksia mahdollistavaa yhteishyväpeliä, joka on tuottanut hyviä tuloksia altruistisen toiminnan kartoittamiseksi. Kappaleessa 3.3 esittelen hyvin käytäntöön sovellettavia sosiaalisen vaihdannan tuloksia, ja lopulta kappaleessa 3.4 tarkastelen neurotaloustieteellistä perustaa altruistiselle ja epäitsekkäälle käyttäytymiselle. Neurotaloustiede on viime aikoina pystynyt luomaan erittäin mielenkiintoisia tuloksia tällä saralla sekä tuottamaan hyvin vartenotettavia teorioita ihmisen perimmäisistä toimintamotiiveista.

3.1. Perinteiset pelien elementit

Taloustieteen parissa on aina elänyt hyvin vahva käsitys itsekkyyshypoteesin puolesta. Tämä käsitys kuitenkin heikkeni 80-luvulla, kun taloustieteilijät alkoivat tehdä pienryhmäkokeita laboratorio-olosuhteissa (Fehr ja Schmidt, 2006). Esittelen tässä kolme yksinkertaista mutta tärkeää ja usklassisille malleille ongelmallista kokeellista peliä: uhkavaatimuspeli (ultimatum game), ulkopuolinen rankaisija -peli (third party punishment game) sekä diktaattoripeli (dictator game). Kaikissa näissä peleissä yhdellä toimijalla on selkeästi dominoiva rooli. Tässä roolissaan hänen on mahdollista ajaa voimakkaasti omaa etuaan. Jos itsekkyysoletus on aina paikkansa

pitävä, dominoiva toimija ajaa aina vain omaa etuaan välittämättä millään tavoin muiden hyödystä. Jos dominoiva toimija taas osoittaa kiinnostusta muiden hyvinvointia kohtaan, voidaan hänen (tai diktaattori- ja ulkopuolinen rankaisija – peleissä ulkopuolisen toimijan) nähdä olevan jossain määrin altruistinen.

3.1.1. Uhkavaatimuspeli

Uhkavaatimuspelissä kahden koehenkilön on pystyttävä jakamaan tietty annettu summa rahaa. Henkilö i toimii ehdottajana. Hän voi tehdä yhden ehdotuksen rahojen jakamisesta. Henkilö j puolestaan toimii vastaajana ja voi joko hyväksyä tai hylätä pelaajan i ehdotuksen. Jos pelaaja j hyväksyy ehdotuksen, jaetaan rahat ehdotetulla tavalla. Jos puolestaan pelaaja j hylkää ehdotuksen, kumpikaan pelaaja ei saa mitään. Jos tässä kontekstissa oletetaan yksilöiden olevan täysin itsekäitä ja rationaalisia ja että henkilö i tietää myös henkilön j olevan itsekäs ja rationaalinen, pelin lopputulokseksi saadaan aina tulos, jossa henkilö i antaa henkilölle j pienimmän mahdollisen rahayksikön ja pitää loput itsellään. Toisaalta henkilö j hyväksyy aina ehdotuksen, koska pienimmänkin rahasumman hyväksyminen edustaa itsekästä ja rationaalista materialistisen hyödyn maksimointia annettuna pelin säännöt.

Fehr ja Schmidt (2006) tiivistävät usean eri tahon tutkimuksen tuloksia. Uhkavaatimuspelin ehdotuksista suuri valtajoukko on 40-50 prosenttia kokonaissummasta pelaajalle j. Toisaalta ehdotukset, jotka tarjoavat pelaajalle j vähemmän kuin 20 prosenttia kokonaissummasta, hylätään todennäköisyydellä 0.4 – 0.6. Lisäksi ehdotuksen hylkäämisen todennäköisyys on vähenevä suhteessa ehdotuksen avokätisyyteen.

Ehdotuksen hylkääminen kertoo Fehrin ja Gächterin (2002) mukaan ilmiöstä nimeltä altruistinen rankaisu (altruistic punishment). Itsekäs pelaaja ei koskaan rankaisisi, koska se on hänelle materialistisessa mielessä epäedullista. Altruistinen henkilö kuitenkin saattaa rangaista toista pelaajaa itsekkyydestä. Tämä rankaisu ei hyödytä rankaisijaa materialistisessa mielessä mutta saattaa vaikuttaa itsekkääseen pelaajaan hänen tulevissa peleissään ja siten auttaa muita pelaajia. Altruistinen rankaisu on oleellinen ilmiö, kun pyritään etsimään hyötyfunktiota, joka voisi rationalisoida kokeista saadun datan. Altruististen preferenssien mallintamiseksi ei siis riitä ainoastaan muiden toimijoiden hyvinvoinnin kasvattaminen vaan tarvitaan selittäjiä itsekkäiden toimijoiden rankaisemiselle.

3.1.2. Ulkopuolinen rankaisija –peli

Tämä peli on erinomainen etenkin edellä esitetyn altruistisen rankaisun tarkastelemiseksi. Pelaaja i toimii ehdottajana ja saa summan S jakaakseen pelaajan j kanssa. Pelaaja j saa pelaajan i tarjoaman summan ilman oikeutta hyväksyä tai hylätä tarjousta. Pelaaja k saa summan $S/2$ ja tarkastelee jakoa. Jaon jälkeen pelaaja k voi määrätä rangaistuksen pelaajalle i . Rangaistuksen määrääminen maksaa pelaajalle k aina yhden yksikön asetettua kolmen yksikön sakkoa vastaan. Pelaaja k voi siis esimerkiksi asettaa pelaajalle i yhdeksän yksikön rangaistuksen maksamalla itse kolme yksikköä. Maksut ja rangaistukset siis menevät ulos pelistä pelin järjestäjälle ja aiheuttavat vain hyödyn laskua pelin yksilöille. Itsekäs pelaaja ei koskaan valitsisi rankaisumahdollisuutta, mutta altruistinen toimija saattaisi määrätä pelaajalle i rangaistuksen normirikkeestä, jos tämä ei ole tarpeeksi avokätinen.

Fehr ja Fischbacher (2004) suorittivat kokeita, joissa kokonaissumman ollessa sata ($S = 100$) pelaaja i sai vain harvoin rangaistuksen jakaessaan pelaajalle j 50 yksikköä tai enemmän. Jos pelaaja i jakoi vähemmän kuin 50 yksikköä, noin 60 prosenttia pelaajista k antoi rangaistuksen, ja rangaistuksen suuruus kasvoi pelaajan i avokätisyyden pienentyessä. Uhkavaatimuspeliin verratessa on kuitenkin huomattava, että henkilön j antama rankaisu ehdotuksen hylkäyksenä oli paljon voimakkaampi kuin ulkopuolisen pelaajan k jakama rangaistus. Fehr ja Schmidt (2006) kuitenkin eivät näe ulkopuolisen rankaisijan jakaman rangaistuksen olevan heikompi sosiaalisten normien ylläpitäjänä, koska ulkopuolisia tarkastelijoita on aina enemmän kuin asianomistajia.

3.1.3. Diktaattoripeli

Diktaattoripelit mittaavat pelaajien puhdasta altruismia. Pelaaja i saa summan S , jonka voi halutessaan jakaa pelaajan j kanssa. Pelaajalla j ei kuitenkaan ole mahdollisuutta rangaista pelaajaa i hänen mahdollisesta itsekkyydestään. Tällöin pelaajan i osoittama avokätisyys voidaan aina tulkita altruistiseksi toiminnaksi. Kappaleessa seitsemän esitetään Andreonin ja Millerin (1998 ja 2002) tarkastelu altruismin rationaalisuudesta diktaattoripelin aineistolla. Kokeissa 22 prosenttia henkilöistä oli täysin itsekkäitä pitäen koko summan S muiden osoittaessa vähintään lievää altruismia toista pelaajaa kohtaan.

3.1.4. Kritiikkiä peleistä

Altruismin taloustieteellisessä kirjallisuudessa on esitetty hyvin vähän kritiikkiä pelejä vastaan. Niihin kuitenkin nähdäkseni liittyy paljon asioita, jotka voidaan tulkita etenkin sosiologian tai antropologian kannoilta hyvin merkittävinä haittatekijöinä.

Kokeet on suoritettu lähes aina laboratorio-olosuhteissa. Tällöin on syytä miettiä, kuinka todellista yksilön käyttäytyminen testitilanteessa on. Pelaaja tietää olevansa asetettuna pelitilanteeseen ja tietää häneen kohdistuvan tiettyjä odotuksia. Esimerkiksi rationaalista toimintaa pidetään yhtenä sosiaalisena normina – haluammehan kaikki toimia järkevästi. On siis luonnollista, että ollessaan tarkkailunalaisena pelaaja pyrkii toimimaan tavanomaista rationaalisemmin. Voi siis olla, että kokeissa havaittu altruismin määrä on vielä todellista tilannetta pienempää. Sinällään tämä kuitenkin vain vahvistaa tarkasteluni tulosta altruismin olemassaolosta.

Kokeiden kautta on myös pyritty tutkimaan kulttuuritaustojen merkitystä tai ylipäänsä luomaan yleisnäkemyksiä keskimääräisestä yksilöstä. Tästä huolimatta testeissä käytetyt henkilöt ovat olleet pääsääntöisesti yliopisto-opiskelijoita. Lienee luonnollista olettaa, että keskimääräisen yliopisto-opiskelijan sosioekonominen asema, ajattelun rakenne sekä arvot poikkeavat muutoin keski-verrosta yksilöstä. Toisaalta toisessa ääri-laidassa talousantropologian tutkimuksissa on käytetty ”köyhiä” tai ns. alkuasukkaita. Tämä on tietysti räikeässä ristiriidassa opiskelijoiden kanssa, ja tulosten vertailukelpoisuus on usein tällöin kyseenalainen. Mikäli toisessa pelissä opiskelijan panoksena on yhden kahvin hinta ja toisessa alkuasukkaalla kuukauden palkka, on selvää, että myös toiminta voi olla erilaista.

Mitä selkeämmin peli näyttää ”rahapeliltä”, sitä haasteellisempia esitetyt ongelmat ovat. Pidemmälle vietyjä, kompleksimpia pelejä ovat esimerkiksi yhteishyväpeli sekä sosiaalisen vaihdannan pelit. Kappaleissa 3.2 sekä 3.3 tarkastellaan näitä pelejä ja syvennetään esitettyjen peruselementtien käyttöä.

3.2. Yhteishyväpeli

Hyvänä esimerkkinä itsekkyyden ja sosiaalisten normien vuorovaikutuksesta yksilön toiminnassa on yhteishyväpeli, joka kokeena selkeästi asettaa nämä käyttäytymistavat vastakkain. Tässä kappaleessa esitellään ensin kohdassa 3.2.1 yhteishyväpelin periaatteet, kohdassa 3.2.2 tarkastellaan empiirisiä tuloksia altruistien ja egoistien stereotyypeistä ja kohdassa 3.2.3 tutkitaan näiden stereotyyppien vaikutusta kokonaisuuteen. Kappaleessa 3.2.4 syvennetään stereotyyppien vaikutus niiden yhteiskunnallisen vaikuttavuuden ja jopa välttämättömyyden tasolle. Tässä tullaan huomaamaan yhteisön jatkuvuuden olevan riippuvainen yksilöiden altruistisista insentiiveistä.

3.2.1. Yhteishyväpelin rakenne

Yhteishyväpelissä jokainen jäsen ryhmästä $N \geq 2$ saa käyttöönsä Y \$. Jokainen ryhmän jäsen voi joko pitää rahat tai investoida ne ryhmän yhteiseen kassaan. Kokeen tekijä kertoo yhteiseen kassaan investoidut rahat tekijällä M , joka on suurempi kuin 1 mutta pienempi kuin N ($1 < M < N$). Kun kaikki ryhmän jäsenet ovat tehneet investointipäätöksensä ja kokeen tekijä on kertonut kassaan kertyneen rahan tekijällä M , yhteinen kassa jaetaan tasan kaikkien ryhmän jäsenien kesken. Jos siis yksilö investoi 5 \$ yhteiseen kassaan, niin koko ryhmä ansaitsee $5xM$ tästä lahjoituksesta, ja jokainen jäsen (mukaanlukien lahjoittaja) ansaitsee $5xM/N$.

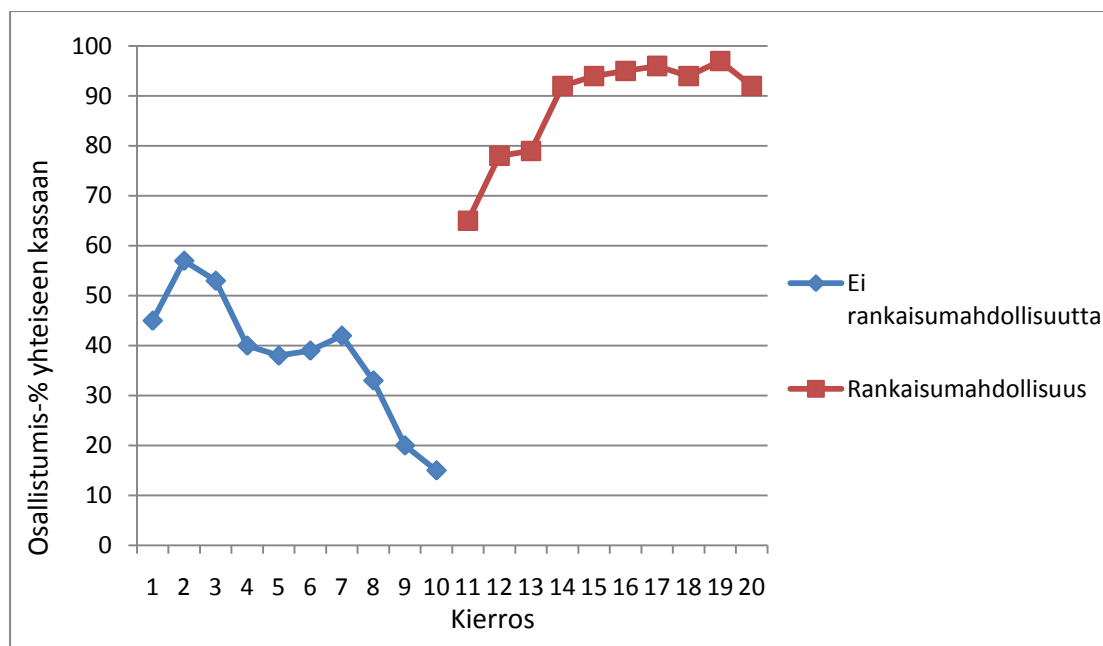
Huomionarvoista on, että lahjoittaja aina huonontaa omaa taloudellista tilannettaan lahjoittamalla, koska $M < N$, mistä seuraa, että alkuperäinen investointi $5 > 5xM/N$. Näin ollen itsekäs yksilö ei koskaan lahjoita yhteishyväpeliin. Toisaalta jos kaikki N jäsentä investoivat koko tulonsa Y , jokainen jäsen ansaitsee $NxYxM/N = YxM$. Esimerkiksi, jos $M = 2$ ja $N = 20$, niin jokainen ryhmän jäsen tuplaa rahansa verrattuna tilanteeseen, jossa kaikki pitävät oman tulonsa. Tosin tilanne tästä huolimatta houkuttaa vapaamatkustajia, koska jokaisesta itse investoidusta dollarista saa takaisin vain kymmenesosan.

Yhteishyväpelille on sovelluskohteita kaikkialla yhteiskunnassa. Tällaisia ovat esimerkiksi turvallisuus, ympäristön saastuminen ja ylipäänsä kaikkien sääntöjen täytäntöönpano, jotka ovat hyväksi yhteisölle mutta aiheuttavat niitä noudattaville yksilöille kustannuksia. Nämä kaikki ilmiöt ovat koetta vastaavan insentiivi-rakenteen varassa. (Fehr ja Gintis, 2007)

Fehr ja Gächter (2000) suorittivat kuvatuunlaisia kokeita tutkimuksissaan ja havaitsivat, että yhden kiinteänä pysyvän ryhmän yhteistyö (eli investointien määrä yhteiseen kassaan) tasoittuu hyvin alhaiselle tasolle useampien kierrosten jälkeen. Heidän kokeessaan oli 6 ryhmää, 10 kierrosta ja muuttujien arvot olivat $N = 4$, $Y = 20$ ja $M = 1,6$. Viimeisellä kierroksella yli puolet yksilöistä ei enää lahjoittanut mitään yhteiseen kassaan ja loputkin vain hyvin vähän. Näiden tulosten valossa on Gintisin mukaan selvästi väärin olettaa, että yhteistyön luoma arvo olisi itsessään riittävä luomaan sosiaalista yhteistyötä, mikä on pitkään toiminut vasta-argumenttina altruistisia preferenssejä vastaan.

Fehr ja Gächter (2000) kuitenkin jatkavat koettaan vielä 10 kierrosta muuttaen sääntöjä hieman. He antoivat osanottajille mahdollisuuden rankaista muita ryhmän jäseniä aina kierroksen lopussa, kun jokaisen ryhmän jäsenen panos oli paljastunut. Jokainen ryhmän jäsen kykeni rankaisemaan toista haluamallaan määrällä ”rangaistuspisteitä”. Tämä alensi sekä rankaisijan että rangaistavan taloudellista tilannetta. Jokainen annettu rangaistuspiste maksoi rankaisijalle yhden dollarin ja rangaistavalle 10% hänen kierroksen jälkeisestä tuloksestaan eli keskimäärin noin kolme dollaria. Tämän pelin lisäpiirteen ei tulisi vaikuttaa pelaajien käyttäytymiseen millään tavalla, mikäli kaikki olisivat itsekkäitä tai yhteisen arvon luominen ei olisi yhdenkään toimijan mielestä mielekästä. Etenkin viimeisillä kierroksilla yhteistyö ja erityisesti rankaiseminen olisi täysin turhaa itsekkäille yksilöille.

Pelin sääntöjen muuttaminen kuitenkin nosti yhteistyöhalukkuutta erittäin dramaattisesti jo ensimmäisellä uusilla säännöillä pelatulla kierroksella ja muutamalla seuraavalla kierroksella se nousi lähes täysimääräiseksi. Kuvassa 1 on havainnollistettuna pelin kulku eri kierroksilla ennen ja jälkeen rangaistusmahdollisuuden käyttöönoton.



Kuva 1. Yhteishyväpelin yhteistyön määrä eri kierroksilla, kun rangaistusmahdollisuutta ei ole ja silloin kun se on mahdollista. (Fehr ja Gächter, 2000)

Fehrin ja Gächterin tutkimuksen viimeiset kymmenen kierrosta siis näyttävät hyvin erilaisilta kuin ensimmäiset kymmenen. Rankaisumahdollisuuden olemassaolo nostaa yhteistyön osuuden lähes täydelliseksi. Vielä viimeiselläkin kierroksella 83% toimijoista lahjoittaa koko summan Y yhteiseen kassaan, ja lopuistakin lähes kaikki lahjoittavat suurimman osan saamastaan summasta. Täysin sama ryhmä ja yksilöt siis saivat aikaan sekä itsekään että altruistisen käyttäytymismallin. Fehr ja Gintis (2007) tulkitsevat kokeiden merkitystä sekä taloustieteen että sosiologian yleisöille painottaen, että oli myös ryhmiä, joissa ensimmäisen kymmenen kierroksen aikana *kukaan ei lahjoittanut mitään* ja viimeisten kymmenen kierroksen aikana *jokainen jäsen lahjoitti kaiken*.

Fehrin ja Gintisin kiinnostuksen kohteena on kaksi täysin erilaista käyttäytymistapaa, joiden perusteella he esittävät tutkimusta motivoivia kysymyksiä: Ensimmäisten kymmenen kierroksen perusteella on selvää, että ihmisissä on paljon itsekkäitä insentiivejä. Kuitenkin on huomattava ensimmäisten muutaman kierroksen noin 50 prosentin osanotto, joka viestii myös altruistisista insentiiveistä, kunnes itsekäs käyttäytymismalli alkaa dominoida. Altruististen insentiivien puolesta puhuu myös viimeiset kymmenen kierrosta ja etenkin voimakas yhteisöllinen käyttäytyminen kierroksella 20. Osallistujat tietävät kokeen päättyvän kierroksella 20. Puhtaasti itsekäs toimija ei koskaan rankaisisi tällä kierroksella, koska se maksaa eikä luo mitään arvoa tulevaisuuteen. Mutta jos kukaan ei rankaisisi kierroksella 20, miksi yksikään rationaalinen ja itsekäs toimija tekisi yhteistyötä tällä kierroksella?

Kaikki edellä esitetyt pelit osoittavat, etteivät yksilöt ainakaan keskimäärin ole *ainoastaan* Homo Economicus –oletuksen mukaisia itsekkäitä ja rationaalisia hyödyn maksimoijia ainakaan perinteisessä instrumentalistisessa mielessä. Jos emme voi olettaa toimijan täydellistä instrumentalistista rationaalisuutta, kuinka voimme tehdä taloustiedettä kuvaamalla yksilöiden hyötyfunktioita? Voiko edellä esitettyjen pelien tuloksia selittää ylipäänsäkään hyötyfunktioiden maksimointina? Näihin kysymyksiin palataan kappaleissa 4 ja 5.

3.2.2. Vahva vastavuoroisuus ja talouden toimijan stereotyypit

Fehr ja Gächter (2002) tulkitsevat aiemmin esitettyjen tutkimustensa tulokset todisteeksi ihmisille luontaisista ja vilpittömistä altruistisista motiiveista – siinä missä toki myös todisteeksi itsekkäistäkin vastaavista. Fehr käyttää useissa julkaisuissaan termiä *altruistinen rangaistus*, jolla hän viittaa toimijalle taloudellisesti ja statuksen kohotuksen kannalta kannattamattomaan vilpittömään ja epäitsekkääseen rankaisemiseen. Tutkimustulokset, jotka osoittavat altruististen rangaistuksien olemassaolon, ovat Fehrin mukaan empiirinen todiste laajemmasta pelkän *antamisen* ylittävästä altruismin käsitteestä, jota hän nimittää vahvaksi vastavuoroisuudeksi. *Vahvan vastavuoroisuuden* (strong reciprocity) käsite on pyrkimys viimein tuoda tarkemmin määritelty vaihtoehto altruismin rinnalle. Vahva vastavuoroisuus vastaa siis näkemystämme altruismista sitomatta sitä kuitenkaan kantilaiseen näkemykseen *kärsimyksen* välttämättömyydestä. Tämän lisäksi se laajentaa käsitteen merkitystä myös rankaisemisen puolelle. Tuleekin muistaa, että vahva vastavuoroisuus ei todellakaan ole vain pitkän tähtäimen valistunutta itsekkyyttä, sillä vahvasti vastavuoroinen toimija tekee yhteistyötä ja rankaisee myös anonyymeissä ja kertaluontoisissa tilanteissa, joissa tämä toiminta ei hyödytä häntä mitenkään *instrumentaalisesti*.

Tutkimustulokset useista erilaisista kokeista – kuten myös yhteishyväpelistä – osoittavat, että merkittävä osa toimijoista todella toimii vahvan vastavuoroisesti. Toisaalta samat kokeet osoittavat, että iso osa toimijoista on puhtaasti itsekkäitä.

Fehr ja Gintis (2007) nostavat vahvan vastavuoroisuuden syiksi ja motiiveiksi kaksi tärkeintä osaa: *vastavuoroinen reiluus* (Rabin, 1993; Falk ja Fichtenbacher, 2006) sekä

epäoikeudenmukaisuuden välttäminen (Fehr ja Schmidt, 1999). Vastavuoroisesti reilua toimijaa motivoi vastata muiden ystävällisiin eleisiin ystävällisesti ja vihamielisiin vihamielisesti. Epäoikeudenmukaisuutta välttävä toimija puolestaan on motivoitunut pyrkimään toimillaan oikeudenmukaisiin lopputuloksiin. Näiden motiivien rinnalla kaikilla toimijoilla kuitenkin luonnollisesti on myös itsekkäitä motiiveja.

Vastavuoroisen reiluuden ja epäoikeudenmukaisuuden välttämisen teorit voivat osaltaan selittää, miksi yhteistyötä on niin vähän rankaisumahdollisuuden puuttuessa ja miksi tuon mahdollisuuden ollessa olemassa yhteistyö on niin vahvaa (Fehr ja Gintis, 2007). Kun rankaisumahdollisuutta ei ole, vahvasti vastavuoroiset toimijat tekevät yhteistyötä, jos uskovat muidenkin tekevän. Uusien kierrosten myötä nämä altruistit kuitenkin huomaavat joukossa olevan vapaamatkustajia, jotka eivät tee yhteistyötä ja näin myös heidän yhteistyönsä laskee hiljalleen kierrosten myötä kunnes lakkaa kokonaan. Huomionarvoista on, että rankaisumahdollisuuden puuttuessa vahvasti vastavuoroiselle ainut tapa rangaista egoistia on lopettaa yhteistyön tekeminen. Näin ollen sekä altruistit että egoistit eivät enää yhteishyväpelin loppukierroksilla lahjoita mitään – joskin heitä ajaa tähän toimintaan eri motiivit.

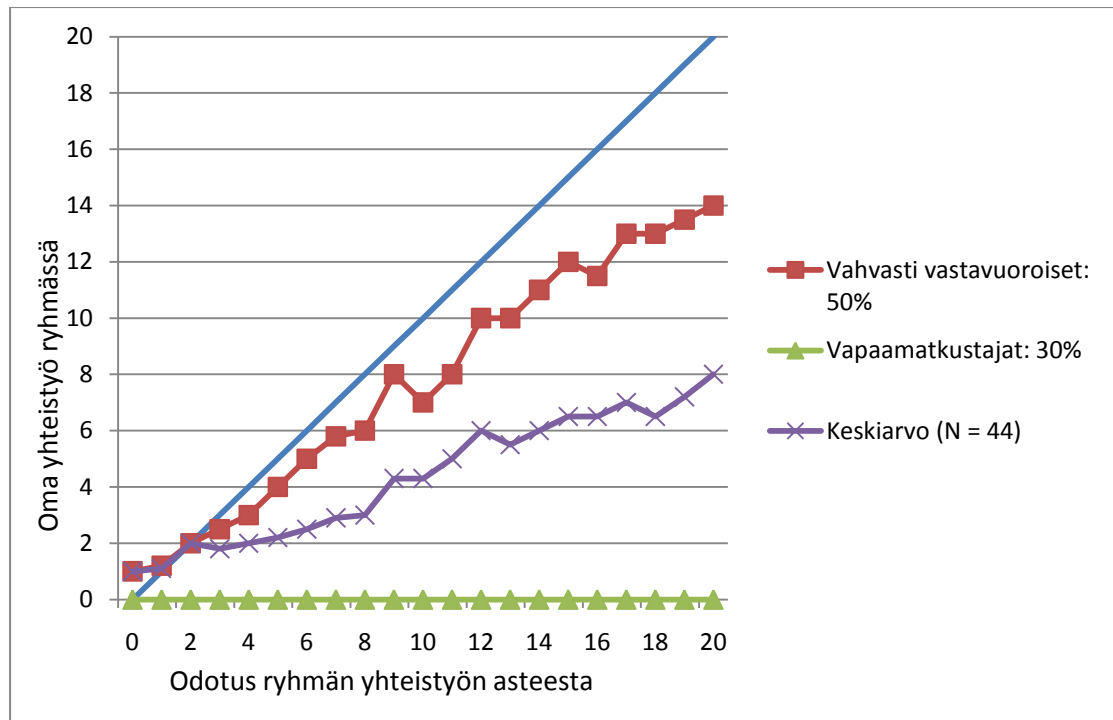
Vastaavasti rankaisumahdollisuuden ollessa olemassa altruistit saavat egoistit lopulta tekemään yhteistyötä. Syyt yhteistyön tekemiseen vastaavasti vaihtelevat. Egoistit tekevät yhteistyötä, koska se on heille edullisempaa kuin rangaistavaksi tuleminen. Vahvasti vastavuoroiset tekevät yhteistyötä, koska egoistitkin näin tekevät ja toisaalta ainoa rankaisukeino ei enää ole yhteistyöstä kieltäytyminen.

Tämä teoria kokeiden empiiristen tulosten selittämiseksi lieneekin Gintistä (2000) ajanut syy tutkia altuismin ja egoismin evolutiivista kehitystä ja tasapainosuhdetta populaatiossa. On myös taloustieteen agregaattien luomisen kannalta tärkeää kyetä määrittämään, mikä on populaatiossa oleva tasapaino altruististen ja egoististen motivaatioiden ajamana toimivien ihmisten välillä. Gintiksen evoluutiomallin tuloksiin palataan kappaleessa 3.2.4.

3.2.3. Altruistiset ja egoistiset stereotyypit yhteishyväpelissä

Kuten aiemmassa kappaleessa kävi hyvin ilmi, arvioidessaan yhteishyväpelien tuloksia Fehr ja Gintis jakavat ihmiset selkeästi altruisteihin ja egoisteihin sen sijaan, että olettaisivat kaikkien toimijoiden toimivan osittain itsekkäiden ja osittain epäitsekkäiden insentiivien motivoimina. Voidaanko heidän ehdottamaansa kahtiajakoa hyväksyä? Jakautuvatko ihmiset todella selkeästi kahteen erilaiseen stereotyyppiin itsekkyyden suhteen?

Fehr ja Gintis tukeutuvat Fischbacherin (Fischbacher et al, 2001) vastaavassa kokeessa – ilman rangaistusmahdollisuutta – saamiin tuloksiin (Kuva 2). Lahjoituspäätöksen lisäksi pelissä pyydettiin yksilöitä arvioimaan kunkin lahjoituspäätöksen jälkeen odotustaan ryhmän ja muiden pelaajien yhteistyön asteesta. Näin Fehr ja Gintis kykenivät vertaamaan yksilön omaa yhteistyön astetta hänen odotukseensa muiden yhteistyöstä.



Kuva 2. Toimijoiden oma yhteistyön aste suhteessa odotettuun muiden yhteistyön asteeseen. Diagonaali kuvaa täydellistä vastavuoroisuutta. (Fischbacher et al, 2001)

Kuvassa 2 neliöt kuvaavat ehdollisesti yhteistyötä tekevän toimijan keskimääräisen käytöksen. Nämä ehdollisesti toimivat nostavat tekemänsä yhteistyön määrää, mikäli uskovat myös muiden ryhmän jäsenten toimivan vastaavasti. Mikäli yksilön käytös kuvautuu diagonaalille, yksilö nostaa yhteistyönsä määrää aina samassa suhteessa kuin uskoo ryhmänkin tekevän. Fischbacherin kokeessa 50% osallistujista toimi tämän ehdollisen yhteistyön mallin mukaisesti. Fehr ja Gintis nimeävät tämän ryhmän kuuluvan vahvasti vastavuoroisiin toimijoihin. Huomionarvoista kyseisen ehdollisen toiminnan osalta on, että sen aggregaatti pysyy hieman ryhmän keskimääräisen odotetun yhteistyön määrän alapuolella eli diagonaalilla $x = y$ alapuolella. Fehrin ja Gintisin mukaan tämä on osoitus siitä, että vahvasti vastavuoroisten toimijoidenkin toiminnassa on itsekkäitä motiiveja. Kuvassa näkyy myös toinen suuri ihmisryhmä,

joka on aina vapaamatkustajina välittämättä toisten panoksesta. Näitä *egoisteja* kokeessa oli 30%. Näiden tulosten perusteella voidaan siis todella olettaa, että suurin osa toimijoista voidaan ryhmitellä kuuluviksi joko egoisteihin tai altruisteihin.

Kuva 2 myös kuvaa hyvin, miksi yhteistyön ylläpitäminen suoran rangaistusmahdollisuuden puuttuessa on mahdotonta. Mikäli kaikki toimijat olettavat kierroksen alussa muiden yhteistyön olevan täydellistä, altruistit lahjoittavat omasta kassastaan suurimman osan mutta eivät kaikkea. Toisaalta joukossa olevat egoistit eivät lahjoita mitään. Näin ollen kierroksen lopussa jokainen toimija saa tiedon, että hänen uskomuksensa toisten käytöksestä olivat väärinä, ja siten jokainen altruisti lahjoittaa seuraavilla kierroksilla entistä vähemmän, kunnes yhteistyö lopulta laskee nolliin.

Fehr ja Gintis (2007) huomauttavat tämän osoittavan myös sen, että yhteiskuntajärjestys voi romahtaa esimerkiksi luonnon mullistusten tai sodan tuloksena, vaikka suurin osa yhteiskunnan jäsenistä olisikin vahvasti vastavuoroisia ja periaatteessa yhteistyöhaluisiakin. Mikäli vapaamatkustajia ei rangaista, myös yhteistyöhaluiset lopettavat yhteistyön tekemisen.

3.2.4. Altruismin yhteiskunnallinen merkitys – evolutiivinen malli

Nyt olemme todenneet, että yhteishyväpeleissä on ollut altruistisia insentiivejä ja altruismia on arkikokemuksen sekä empiiristen todisteiden valossa olemassa. Jotta voimme vielä todeta altruismin todella olevan osa ihmisen luonnetta, tarkastelen tässä kappaleessa mallia, joka osoittaa altruismin olevan kaikkea muuta kuin triviaali ilmiö

yhteisössä. Seuraava malli esittää altruististen insentiivien yhteiskunnallisten vaikutusten olevan merkittäviä yhteiskunnan kokonaisyhyvinvoinnin kannalta.

Yhteistyön jatkuvuus ja olemassaolo yhteisössä on herkkä asia. Siihen vaikuttavat jäsenistön insentiivit, yhteisön keston todennäköisyys ja muu ympäristö. Näistä syistä yhteistyön jatkuvuus on pitkän historiamme ja ihmisrodun evoluution kannalta mielenkiintoinen tarkasteltava. Gintis (2000) on pyrkinyt luomaan ja esittämään matemaattisen mallin, joka kuvaisi altruististen ja egoististen toimijoiden osuutta yhteisössä, jotta yhteistyö on vielä mahdollista ja kestävä. Tämä malli on mielenkiintoinen myös siksi, että se kuvaa keskeisimmät tekijät altruistisen yksilön taloudellisessa päätöksenteossa ja antaa siten myös näkökulmaa hyötyfunktioiden mallintamiseen. Tässä pyrimme lähinnä tarkastelemaan Gintisin mallin johtopäätöksiä ja niiden merkitystä.

Gintisin tarkastelu pohjautuu yhteishyväpeliin, jonka jo aiemmin totesimme hyväksi ja monipuoliseksi peliksi yksilön insentiivien tarkastelussa. Gintisin mallissa jokaisella kierroksella ja jokaisella ryhmän henkilöllä on mahdollisuus uhraamalla $c > 0$ kerryttää ryhmän yhteistä hyvää $b > c$ verran. Jos kaikki toimivat yhdessä, jokainen saa nettotuloksen $b - c > 0$. Jokaisen kierroksen päätteeksi ryhmässä hyljeksitään yksilöitä, jotka eivät lahjoittaneet c :tä, koska ryhmän vahvasti vastavuoroisilla on halu rankaista vapaamatkustajia. Ryhmä pysyy kierroksen jälkeen koossa todennäköisyydellä δ . Gintis lisäksi olettaa, että on olemassa niin sanotuja kriittisiä tiloja, jotka ilmenevät kierroksella todennäköisyydellä p . Kriittisessä tilanteessa todennäköisyys yhteisön koossa pysymiseen δ_* on alhaisempi kuin normaalilla kierroksella δ^* .

Ensin Gintis tarkastelee egoistisen yksilön päätöksentekoa yhteistyön suhteen laskemalla egoistisen yksilön odotusarvoa. Odotusarvo päätökselle toimia yhteistyössä muodostuu tällöin yksilön nettotuloksesta $b - c$ sekä mahdollisuudesta että peli jatkuu todennäköisyydellä δ ja tuottaa jälleen saman lopputuloksen π . Näin kaavasta $\pi = b - c + \delta\pi$ Gintis johtaa ensimmäisen teoreemansa, jonka mukaan egoistiset yksilöt toimivat yhteistyössä, jos ja vain jos

Teoreema 1. $\frac{c}{b} \leq \delta.$

Mikäli siis yhteistyön tekemisen odotusarvoiset hyödyt ovat korkeammat kuin ensimmäisen kierroksen rahastaminen ja yhteisöstä pois sulkeutuminen, egoistinenkin yksilö tekee yhteistyötä. Epävarmuustekijän δ olemassaolo kuitenkin tarkoittaa sitä, että mikäli ryhmä hajoaa suurella todennäköisyydellä, itsekkäät toimijat eivät tule tekemään yhteistyötä – oli yhteistyön tekemisen hyödyt sitten kuinka suuret tahansa. Teoreeman 1 mukaan puhtaasti egoistisen yhteisön yhteistyö vaatii aina uhrauksen suhteen sen hyötyyn olevan pienempi kuin todennäköisyys yhteisön koossa pysymiseen, koska vain tällöin odotusarvo yhteistyöstä on positiivinen.

Gintis (2000) kuitenkin huomauttaa, että *Homo sapiensin* evolutiivisessa historiassa on ollut paljon suuria muutoksia populaatioiden koossa, mikä tarkoittaa, että sosiaalisia kriisejäkin on ollut. Koska yhteistyötä ihmispopulaatioissa on ollut halki sen historian, Gintis ehdottaa toista lähestymistapaa vahvasti vastavuoroisen toimijan toiminnan kartoittamiselle.

Toisin kuin teoreeman 1 egoistiset toimijat, vahvasti vastavuoroinen ihminen tekee yhteistyötä ja rankaisee vapaamatkustajia välittämättä arvosta δ , vaikka tulevaisuus

näyttäisikin huonolta. Tämä todettiin jo aiemmin käsitellyssä yhteishyväpelissä, jossa vahvasti vastavuoroiset toimijat tekivät yhteistyötä hyvin jopa viimeisellä kierroksella.

Vahvasti vastavuoroiset toimijat ovat altruisteja, koska he lisäävät toisten (lähipiiriin kuulumattomien) yhteisönsä toimijoiden hyvinvointia, vaikka se aiheuttaisi heille kustannuksia. Jos vahva vastavuoroisuus on evoluution tulosta, Gintisin mukaan on vahvasti vastavuoroisten yksilöiden oltava selvä hyöty yhteisölleen. Lisäksi ryhmän hyötyjen on ylitettävä yksilön kustannukset.

Valittuaan yhtälön, jonka uskoo toteuttavan edellä mainitut hyödyt ja kustannukset, Gintis muotoilee kolme teoremaa lisää: yhden liittyen vallitseviin olosuhteisiin (teoreema 3) ja kaksi liittyen stereotyyppien osuuksiin yhteisössä (teoreemat 2 ja 4). Meitä kiinnostaa nyt lähinnä jälkimmäinen kysymys.

Gintisin teoreema 2 kuvaa eri tekijöiden vaikutuksia tarvittavaan altruististen toimijoiden osuuteen yhteisön populaatiosta, jotta yhteistyö säilyy.

Teoreema 2. *Yhteistyöhön vaadittujen vahvasti vastavuoroisten osuutta f*

1. *vähentää huonon tilan toteutumisen todennäköisyyden p pieneneminen,*
2. *selviytymisen todennäköisyyden δ_* kasvaminen huonolla kierroksella ja/tai*
3. *vahvasti vastavuoroisten aiheuttaman rangaistuksen määrän kasvaminen.*

Näillä ominaisuuksilla on melko suoraviivaiset tulkinnat. Huonon tilan toteutumisen todennäköisyyden p pieneneminen nostaa yksilön odotusarvoa hänen henkilökohtaisesta tulostaan (egoistilla $b - c + \delta\pi$), koska yhteisön jatkuvuuden todennäköisyys δ tulevilla kierroksilla on suurempi (termeillä p ja δ todettiin jo

aiemmin olevan suora positiivinen korrelaatio). Tämä alentaa hyötyä egoistiselle toiminnalle $b - \delta\pi$ huonossa tilassa. Egoistisen toiminnan ollessa kannattamattomampaa on luonnollista, että entistä pienemmällä rangaistuksella on myös mahdollista saada yhteistyön odotusarvo egoistista toimintaa suuremmaksi. Tällöin entistä pienempi määrä vahvasti vastavuoroisia kykenee ylläpitämään yhteistyötä. Selviytymisen todennäköisyyden δ_* kasvaminen nostaa odotusarvoa π ja alentaa egoismin hyötyä $b - \delta\pi$ aivan samoin päättelyin ja tuloksin.

Vahvasti vastavuoroisten yksilöiden aiheuttaman rangaistuksen voimakkuuden kasvu on teoreeman väitteistä selvin. Kun altruisti kykenee samalla vaivalla aiheuttamaan egoistille entistä voimakkaamman rangaistuksen, ei yhteisö tarvitse yhtä montaa altruistia yhteistyön ylläpitämiseksi. Siinä missä eläinten välisessä välienselvittelyssä lopputulos vaati yleensä ison kustannuksen myös voittajalta, ihmisyhteisöissä pienikin määrä hyökkääjiä kykenee voittamaan valtavimmankin yksittäisen vihollisen pienellä kustannuksella käyttäen koordinaatiota, viekkautta sekä tappavia aseita (Gintis, 2000). Vastaavasti nyky-yhteiskunnassa esimerkiksi tyytymätön kuluttaja kykenee äärimmäisen pienellä vaivalla aiheuttamaan myyjälle isoja harmeja internetin yhteisöpalveluja hyödyntämällä.

Kun altruistit rankaisevat itsekkäitä yksilöitä, altruistit kärsivät kulun c , jota Gintis kutsuu myös *valvontakuluksi*. Rangaistuksia jaetaan vain huonoilla kierroksilla, kun yhteistyötä ei tehdä. Tällöin vahvasti vastavuoroiset kärsivät kulun c vain todennäköisyydellä p . Odotettu kulu rankaisemisesta on siis pc

Teoreema 4. *Egoistiset toimijat kykenevät aina tunkeutumaan vahvasti vastavuoroisten toimijoiden joukkoon. Samoin, kun egoistien rankaisemisen kustannus c on riittävän pieni, pystyy pienikin joukko vahvasti vastavuoroisia tunkeutumaan egoististen toimijoiden joukkoon.*

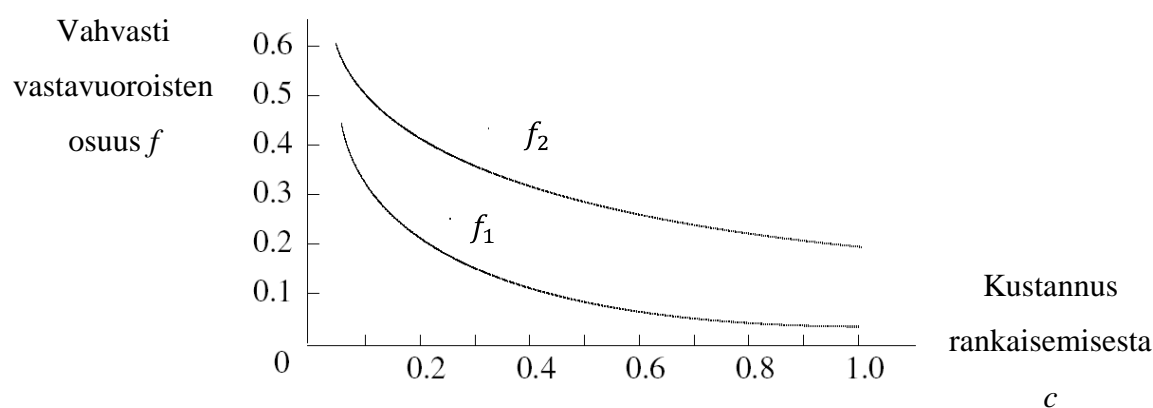
Gintis myös olettaa, että yhteisöjen syntyessä egoistien ja altruistien osuus yhteisössä on verrannollinen niiden osuuteen koko populaatiossa. Tällöin jokaisessa uudessa yhteisössä on sekä altruisteja että egoisteja. Mikäli yhteisö toimii hyvin kaikkien jäsentensä parhaaksi ja egoismista langetetut rangaistukset ovat riittäviä, egoistilla on insentiivi toimia yhteistyössä muiden kanssa. Tällöin egoisti ei sinällään poikkea altruistisesta toimijasta ja kykenee toimimaan osana valtaosin altruististakin yhteisöä tulematta karkoitetuksi tai hyljeksityksi. Vastaavasti egoistien rankaisemisen kustannuksen ollessa pieni pystyy pienikin altruistien joukko luomaan riittävän insentiivin yhteistoiminnalle egoistisessa yhteisössä, kuten myös teoreeman 2 jälkeen todettiin.

Gintisin teoreemien 2 ja 4 mukaan vahvasti vastavuoroisten aiheuttama kulu rangaistuksista egoisteille sekä tästä aiheutuva kustannus vahvasti vastavuoroisille toimijoille ovat kantavia voimia kyseessä olevien ominaisuuksien tasapainon olemassaololle ja sen ylläpidettävyydelle.

Mallin mukaan egoistit saattavat tietyissä olosuhteissa kyetä toimillaan lakkauttamaan kaiken altruistisen toiminnan yhteisöstä, mutta tällöin yhteisö myös hajoaa epävarman tulevaisuuden tullessa kyseeseen. Toisaalta vahvasti vastavuoroiset eivät missään tapauksessa kykene ajamaan egoistisia toimijoita sukupuuttoon. (Gintis, 2000)

Mallin muodostettuaan Gintis simuloi sitä paljastaakseen vahvan vastavuoroisuuden tasapaino-osuuden f populaatiossa. Kuvan 3 alempi käyrä näyttää vahvasti vastavuoroisten tasapaino-osuuden rangaistuskustannuksen c funktiona välillä $[0,05; 1]$. Tälle käyrälle Gintis olettaa 40 jäsenen ryhmän, jonka selviytymisen todennäköisyys hyvässä tilassa δ^* on 0,95 ja selviytymisen todennäköisyys kriittisessä tilassa δ_* on 0,10. Huonon kierroksen todennäköisyydeksi p Gintis olettaa 0,10 ja tällaisella kierroksella yhteistyön tuottamiseksi oletetaan tarvittavan vahvasti vastavuoroisten osuudeksi $3/8$.

Ylempi käyrä näyttää saman suhteen kun ryhmässä on kahdeksan jäsentä. Käyrä on simuloitu erilaisilla oletusarvoilla, joiden Gintis uskoo kuvaavan paremmin esimerkiksi tiivistä perheyhteisöä, jossa vahvan vastavuoroisuuden ominaisuus on selkeästi periytyvä perheiden sisällä. Sen tarkoituksena on lähinnä osoittaa, että altruismin esiintyminen voi olla huomattavasti suurempaakin etenkin kalliin rankaisun tilanteissa.



Kuva 3. Tasapaino altruistien osuuden ja rangaistuskustannuksen välillä yhteisön ylläpitämiseksi (Gintis, 2000).

Kuvan 3 tulokset toistavat jo aiemmin tehtyjä päätelmiä Gintisin mallista. Mikäli altruististen toimijoiden osuus yhteisössä on pieni, heidän on kannettava suurempi kustannus egoistisen toiminnan rankaisemisesta yhteistyön ylläpitämiseksi. Toisaalta mikäli altruistisia toimijoita on paljon, rankaisemisen kustannukset muodostuvat pienemmiksi.

Gintis on luonut mallin, joka kykenee selittämään yhteistyön olemassaolon ja säilyvyyden evolutiivisessa kehityksessä. Vastavuoroisen altruismin eli heikon vastavuoroisuuden malli, joka on laajemmin taloustieteenkin kentässä hyväksytty, ei kykene vastaavasti selittämään pitkän ihmisevoluution aikana ylläpidettyä korkeahkon yhteistyön astetta. Tämä johtuu siitä, että yhteiskuntien voidaan olettaa kokevan suuria sukupuuttollakin uhkaavia tapahtumia, joiden aikana vastavuoroinen altruismi ei kykene motivoimaan egoistisia toimijoita riittävästi. Mikäli vahvasti vastavuoroisten osuus on riittävän suuri, jopa egoistit voidaan saada tekemään yhteistyötä vaikeissakin tilanteissa. Tämä pienentää sukupuuton todennäköisyyttä.

Tämä malli antaa tukea kaikille ja tukeutuu kaikkiin edellä tarkasteltuihin kokeisiin. Se pohjautuu ajatuksiin altruististen ja egoististen stereotyyppien olemassaolosta. Tämän lisäksi se kuvaa altruismin välttämättömänä yksilön ja yhteisön ominaisuutena siinä, missä Homo Economicus on tulkinnut sen lähes paheeksi.

3.3. Sosiaalinen vaihdanta ja vahva vastavuoroisuus

Edellä osoitettiin, kuinka erilaiset sosiaaliset ja yhteiskunnalliset mallit, kuten eri tasoisten rangaistusmahdollisuuksien olemassaolo, voivat luoda erilaisia käyttäytymismallien agregaatteja. Jo tämä sinällään osoittaa, että taloustiedettä ei voida perustaa pelkälle yksilön käyttäytymistä ja preferenssejä kuvaavalle teorialle. Kuten kokeissa osoitettiin, samat henkilöt voivat saada aikaiseksi täysin erilaisia käyttäytymismalleja eri olosuhteissa. Fehrin ja Gintisin (2007) mukaan heidän teoriansa ei kuitenkaan sivuuta yksilöllisiä ominaisuuksia, vaan pikemminkin auttaa kahden tärkeän yksilöllisen ominaisuuden (itsekkyyden ja vahvan vastavuoroisuuden) tunnistamisessa empiirisen tutkimusaineiston perusteella. Etenkin talouden toimijan altruististen insentiivien tunnistamisessa Fehr ja Gintis kokeillaan onnistuvat. Mikäli perinteinen, egoistinen Homo Economicus -oletus pitäisi paikkansa eikä toimijoilla olisi minkäänlaisia epäitsekkäitä insentiivejä, molemmat pelin säännöt yhteishyväpelissä tuottaisivat täysin samanlaisen tuloksen.

Fehr ja Gintis soveltavat tuloksiaan myös sosiaaliseen vaihdantaan. Vaikka vaihdannan teoriat ovat usein toimivia myös perinteisen Homo Economicus -oletuksen valossa, myös vaihdannassa on Aspekteja, jotka vaativat altruistisia insentiivejä. Tieteenteoreettisesta näkökulmasta ei lieneäkään hyväksyttävää, että vaihdannan teorioissa tyydyttäisiin rationaalisen egoistin oletukseen vain, koska se sattuu toimimaan. Kuinka esimerkiksi sosiaalinen vaihdanta saa alkunsa puhtaasti egoististen ja toisiaan egoistisina pitävien toimijoiden kesken? Kuinka sosiaalinen normisto syntyy, mikäli sille sinällään ei ole vahvasti vastavuoroisia käyttäjiä, jotka sitä alkaisivat ylläpitämään?

3.3.1. Sosiaalisen vaihdannan kokeet – työmarkkinat

Fehr ja Gintis ottavat tulkintansa kohteeksi vaihdantakokeen (Fehr, Kirchsteiger ja Riedl, 1993), joka perustuu työnantajan palkanmaksuun sekä työntekijän työpanoksen valintaan. Tässä kokeessa yksi osallistuja ottaa työnantajan roolin ja tarjoaa palkkaa W toiselle osallistujalle, joka työntekijän roolissa tuottaa työnantajalle panoksen E vastineeksi palkasta W . Ensin työnantaja tarjoaa sitovaa palkkaa W työntekijälle, minkä jälkeen työntekijä saatuaan tietoonsa tarjotun palkan vastaa tarjoamalla työpanosta E . Työnantaja ei enää tässä vaiheessa voi vaikuttaa palkkaan W , oli tarjottu panos E mieluisa tai ei.

Työnantajan rahallinen hyöty π voidaan ilmaista kaavana $\pi = AxE - W$, jossa A kuvaa työn tuottavuutta (esimerkiksi tehokkuuden tunnuslukua jalostusarvo/henkilöstökulut) ja on kokeen tekijän päättämä vakio. Tällöin työnantajan hyöty kasvaa, kun työpanos E kasvaa, ja laskee, mikäli palkka W nousee. Työntekijän hyöty U puolestaan saadaan kaavasta $U = W - C(E)$, missä $C(E)$ kuvaa panoksen kustannusta työntekijälle. Koska kokeessa korkeampi panos E tarkoittaa myös korkeampaa kustannusta C työntekijälle, korkeampi työpanos laskee hyötyä U .

Jotta saataisiin aikaan yksinkertaista sosiaalisen vaihdannan tilannetta vastaavat puitteet, joissa työntekijä on täysin vapaa itse valitsemaan haluamansa työpanoksen E ilman minkäänlaisia ulkoisia rajoituksia esimerkiksi sopimusten tai sopimusrikkeiden rangaistusten suhteen, pidetään koe kertaluontoisena aina kahden pelaajan välillä.

Minkäläistä käyttäytymistä pitäisi odottaa? Mikäli kaikki toimijat noudattavat egoistista Homo Economicus -hypoteesia, työntekijä tarjoaa aina pienimmän

mahdollisen panoksen E , koska se maksimoi hänen hyötynsä U . Työnantaja puolestaan valitsisi aina pienimmän mahdollisen W , koska se maksimoi hänen tulonsa π . Fehrin ja Gintisin mallin mukaan on kuitenkin mahdollista, että työnantajalla on motiivi maksaa reilua palkkaa, mikäli joukossa on riittävästi vahvasti vastavuoroisia toimijoita. Tämä johtuu siitä, että vahvasti vastavuoroisilla toimijoilla on tapana vastata ystävällisiin tekoihin eli reiluun palkkaan reilulla työpanoksella. Tällöin työnantajalla on insentiivi reilun palkan maksamiseen riippumatta hänen egoistisista motiiveistaan.

Fehrin, Kirchsteigerin ja Riedlin (1993) kokeissa todella kävi kuten Fehr ja Gintis olettavat: Työntekijät valitsivat keskimäärin korkeamman panoksen, kun työnantajat maksoivat korkeampaa palkkaa. Yhteistyö, kaikkien ulkoisten kannustimien ja rajoitteiden puutteesta riippumatta, on osoitus vahvasti vastavuoroisten toimijoiden olemassaolosta pelissä. Vahvasti vastavuoroisten yksilöiden vaikutus pelin tuloksiin osoittaa Fehrin ja Gintisin (2007) mukaan, että myös osa sosiaalisen vaihdannan kannustimista johtuu ihmisissä olevista sisäisistä sosiaalisista arvoista eli jonkinasteisesta altruismista. Näin ollen altruististen aspektien tuominen mukaan myös vaihdannan tarkasteluihin näyttää perustellulta.

3.3.2. Sosiaalisen vaihdannan seuraukset

Perinteisen teorian mukaan palkka työmarkkinoilla määräytyy kysynnän ja tarjonnan perusteella. Edellä kuvattujen tulosten mukaan palkkaan kuitenkin todellisuudessa saattaa vaikuttaa myös sosiaalisen vaihdannan aspekti. Suuri vahvasti vastavuoroisten toimijoiden osuus työmarkkinoilla antaa työnantajille insentiivin reilun palkan

maksamiseen markkinoiden tilasta riippumatta, koska se saattaa vaikuttaa heidän saamansa työpanoksen laatuun ja määrään.

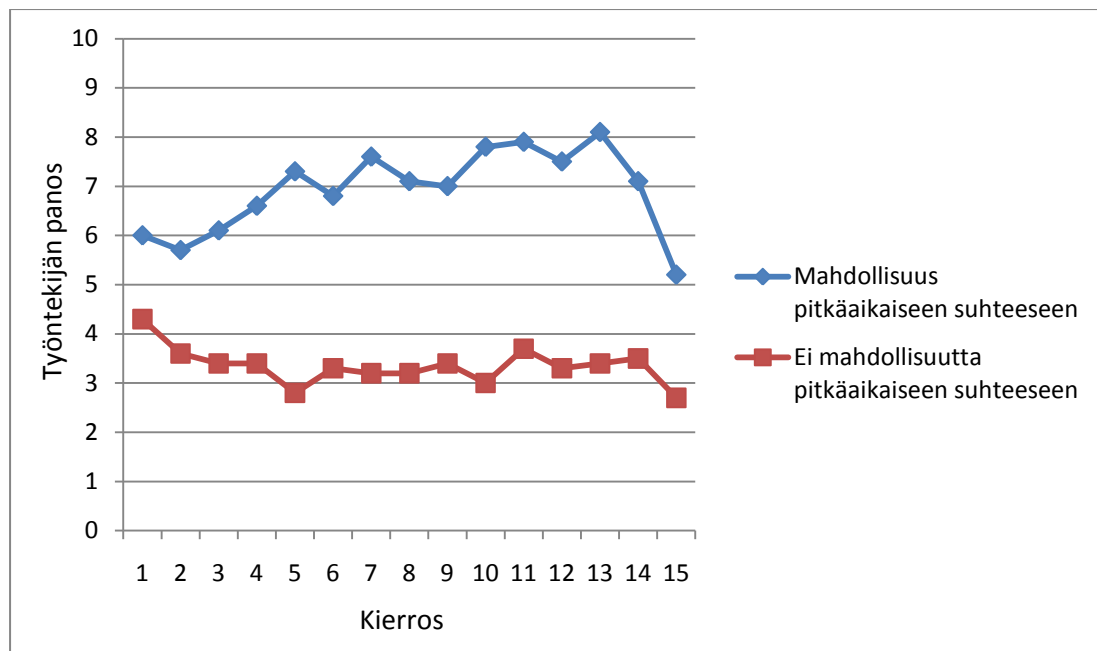
Fehrin ja Falkin (1999) tutkimuksissa toteutettiin vastaava koe kuin edellisessä kappaleessa sillä erolla, että työntekijöitä oli enemmän kuin työpaikkoja – työmarkkinoilla oli siis ylitarjontaa. Työntekijöiden piti kilpailla työpaikoista, minkä kysynnän ja tarjonnan lakien mukaan tulisi laskea palkkatasoa. Fehr ja Falk kuitenkin havaitsivat, että vahvasti vastavuoroisten toimijoiden takia työnantajilla oli insentiivi maksaa reilua palkkaa, ja siten palkkataso ei muuttunut aiemmasta tasapainotilasta mihinkään. Tämän tutkimuksen valossa voitaisiin siis tulkita, että hinnan muodostukseen vaikuttavat paitsi kysyntä ja tarjonta myös sosiaaliset normit reiludesta sekä niitä noudattavien ja ne sisäistäneiden osuus työmarkkinoilla. Altruistisilla insentiiveillä on suora vaikutus hyvin laajalti taloustieteen teoriaan sekä sovellutuksiin, ja niiden perinteinen huomiotta jättäminen saattaa johtaa edellä kuvatuun kaltaisiin vääriin tulkintoihin todellisuudesta.

Tähän mennessä tarkastellut kokeet eivät kuitenkaan ole huomioineet pitkäkestoisiin suhteisiin pyrkimisen mahdollisuuksia. Tämän mahdollisuuden poissulkeminen onkin altruististen insentiivien tarkastelun kannalta äärimmäisen tärkeää. On kuitenkin muistettava, että on monia suhteita ja sopimuksia, joissa lain kaltaista kolmannen osapuolen *vahvistajaa* ei ole, kuten esimerkiksi kansainvälinen kauppa. Tällöin ainut tapa pyrkiä kestävään vaihdantaan on etsiä ja luoda pitkäkestoisia suhteita.

Brown ja kumppanit (2004) toistivat aiempaa vastaavan kokeen tuoden siihen kuitenkin myös pitkäkestoisten suhteiden mahdollisuuden. Koe kesti 15 kierrosta sekä kertaluontoisena että suhteiden luonnin mahdollistavana. Työntekijöistä oli jatkuva

ylitarjonta, ja osa työntekijöistä oli aina työttöminä. Kun työntekijä, jolle oli tehty työtarjous, hyväksyi ehdotetun palkan, hän sai valita työpanoksensa määrän välillä $[0, 10]$. Pitkäkestoiset suhteet mahdollistavan osuuden aikana työntekijät olivat identifioitavissa heille nimettyjen numeroiden perusteella siinä missä kertaluonteisten kierrosten aikana edellisen kierroksen työntekijää ei kyetty identifioimaan. Näin ollen suhdetoiminnan mahdollistavilla kierroksilla työnantaja pystyi aina kierroksen jälkeen palkkaamaan saman työntekijän uudestaan ja tekemään tälle uuden palkkaehdotuksen.

Kuva 4 esittää kokeen keskimääräisen työpanoksen tuloksia suhteessa pelattuihin kierroksiin. Kertaluontoisessa pelissä yhteistyö on huomattavasti alhaisempaa kuin pidemmät suhteet sallivassa pelissä, jossa suhteiden muodostuminen luo insentiivin myös itsekkäälle toimijalle. Jos pitkäkestoisten suhteiden luominen ei ole mahdollista, itsekkäät toimijat valitsevat aina minimaalisen työpanoksen tason. Vain vahvasti vastavuoroiset toimijat tarjoavat ei-minimaalisen työpanoksen, mikäli heille tarjotaan reilua palkkaa. Pitkäkestoiset suhteet salliva peli kuitenkin mahdollistaa myös itsekkäiden toimijoiden rankaisemisen altruistin jättäessä tämän uudestaan palkkaamatta, mikäli työntekijän panos ei ole riittävässä suhteessa tarjottuun palkkaan. Työnantajat puolestaan tiedostavat tämän mahdollisuuden sekä vahvasti vastavuoroisten työntekijöiden olemassaolon, joten heillä on insentiivi maksaa reilua palkkaa – jopa viimeisellä kierroksella. Näistä syistä nyt myös itsekkäällä toimijalla on vahva insentiivi tarjota reilu työpanos vastineeksi palkasta. (Fehr ja Gintis, 2007)



Kuva 4. Suhteiden jatkuvuuden vaikutus sosiaalisen vaihdannan tuloksiin. (Fehr ja Gintis, 2007)

Fehr ja Gintis käyttävät kuvaa 3 puhtaasti tehdäkseen johtopäätöksiä *työntekijöiden* käytöksestä eri pelien välillä. On kuitenkin muistettava, että valittuihin työpanoksiin vaikutti aina työnantajan esittämä palkka, jonka suuruuteen puolestaan vaikuttaa, kuten aiemmissa vastaavissa peleissä todettiin, työnantajan odotukset työntekijöiden toiminnasta vallitsevissa olosuhteissa. Näin ollen kuva 4 ei esitä ainoastaan työntekijöiden insentiivejä vaan ylipäänsä pelissä vallinnutta yhteistyön määrää. Esittämäni ongelman ratkaisujen kannalta tämä kuitenkin on yhdentekevää.

3.4. Altruismin neurotaloustieteellinen perusta

Fehr ja Schmidt (2006) esittelevät artikkelissaan neurotaloustieteen viimeisimpiä saavutuksia aivokuvantamisen hyödyntämisestä behavioristisissa kokeissa. Heidän mukaansa striatum-niminen aivojen osa on aivojen neuroverkkojen avain tarkasteltaessa hedonistisia kokemuksia. Useiden heidän esittämiensä

tutkimustulosten mukaan altruistiset toimet – jopa altruistinen rankaisu – herättää striatumissa aktiivisuutta, joka viittaa hedonistiseen kokemukseen. Huomionarvoista on myös, että tämä tulos on suora hyökkäys altruismin kantilaista määritelmää kohtaan. Mikäli teko tuottaa hedonistista nautintoa, sen taustalla on väistämättä muutakin kuin eettisiä velvollisuuksia.

Nämä tulokset neurotieteen mukaan osoittavat, etteivät ihmiset toimi altruistisesti maksimoidakseen materiaalista hyötyään, vaikka se olisi odotettavissa vasta tulevaisuudessa. Perinteisestihän taloustiede on ehdottanut altruismin kumpuavan odotetusta tulevaisuuden hyödystä vastavuoroisuuden kautta. Sen sijaan altruistinen käytös itse asiassa herättää toimijassa välittömästi tyytyväisyyden kokemuksia riippumatta materialistisista vaikutuksista.

Nämä tulokset puoltavat näkemystä altruistisista preferensseistä. Yksilöiden on nyt täysin rationaalista toimia altruistisesti, sillä heidän tyytyväisyytensä – eli laajassa merkityksessä hyöty – kasvaa altruistisesta toiminnasta. Tämä kuitenkin viittaa nimenomaan hyödyn kokemiseen niin epäinstrumentalistisena kuin materialistisenakin etuna toisin kuin taloustieteen dualistinen näkemys ja Homo Economicus -oletus on asiaa kuvannut. Tällöin rationaalisuus tulisi, materiaallisen hyödyn maksimoinnin sijaan, nähdä vain kykynä valita korkein indifferenssikäyrä hyötyfunktion ollessa konkaavi. Nämä todisteet mahdollistavat hyötyfunktioit semanttisella tasolla altruismin ollessa selitettävissä niihin kytkettynä. Seuraavan kappaleen tarkastelu pohjautuu nimenomaan tämänkaltaiseen laajempaan käsitykseen rationaalisuudesta.

Kappaleen 3 perusteella kuitenkin uskallan nyt todeta, että vastaus ensimmäiseen tutkimuskysymykseeni on saavutettu. *Altruistista käyttäytymistä todella on olemassa.*

4. Onko altruismi rationaalista?

Edellä on esitetty lukuisia todisteita altruististen insentiivien olemassaolosta, mikä puhuu vahvasti taloustieteen perinteistä Homo Economicus -oletusta vastaan. Tarkoittaako tämä, että ihmiset eivät olekaan rationaalisia, vai voimmeko nähdä myös altruistisessa toiminnassa rationaalisia piirteitä? Onko rationaalisuus siis vain taloudelliseen ja mitattavaan liittyvä termi vai voidaanko se liittää myös pehmeisiin suureisiin?

4.1. Altruistisen hyötyfunktion mahdollisuus

Andreoni ja Miller (1998 ja 2002) esittävät, kuinka diktaattoripelissä esiintyvä altruismi on yhdenmukainen kuluttajanteorian kolmen valinnan aksiooman kanssa. Heidän mukaansa ”rationaalisuus” on nimenomaan käyttäytymistä, joka on yhtenäistä näiden aksioomien kanssa, eli käyttäytymistä, jota voidaan mallintaa konvekseilla preferensseillä. Tällöin on myös mahdollista kuvata hyötyfunktio, joka tuottaa järkeviä indifferenssikäyriä.

Olkoon x_i monetaarinen tulos, jonka diktaattori valitsee itselleen, ja x_j tulos, jonka diktaattori antaa toiselle. Tällöin kysymyksenasettelu on: Voidaanko kokeiden käyttäytymistä kuvata kvasi-konkaavilla hyötyfunktioilla, joka on muotoa $U_i=(x_i,x_j)$?

Jos siis tällainen ”oikein” käyttäytyvä hyötyfunktio on voinut tuottaa peleistä saadun datan, on sen täytettävä seuraavat kolme aksioomaa ja siten myös yleistetyn paljastettujen preferenssien aksiooman (General Axiom of Revealed Preferences).

Paljastettujen preferenssien kuluttajan teorian aksioomat ja määritteet voidaan yksinkertaistaa seuraavasti (Andreoni ja Miller, 2002):

Paljastettujen preferenssien heikko aksiooma WARP (Weak Axiom of Revealed Preference): Jos allokatio X on suoraan paljastettu preferoiduksi allokatioon Y , ei Y voi olla suoraan paljastettu preferoiduksi allokatioon X .

Määritelmä: Suoraan paljastettu preferoiduksi (Directly Revealed Preferred). Allokatio X on suoraan paljastettu preferoiduksi allokatioon Y , jos Y oli valittavissa, kun X valittiin.

WARP on Andreonin ja Millerin mukaan pakollinen, muttei riittävä, ehto yhtenäisyydelle hyötyteorian kanssa.

Paljastettujen preferenssien vahva aksiooma SARP (Strong Axiom of Revealed Preferences): Jos allokatio X on paljastettu preferoiduksi allokatioon Y , ei Y koskaan ole paljastettu preferoiduksi allokatioon X .

Määritelmä: Paljastettu preferoiduksi (Revealed Preferred). Jos allokatio A on suoraan paljastettu preferoiduksi B :hen ja B on suoraan paljastettu preferoiduksi C :hen, A on paljastettu preferoiduksi C :hen.

Jos preferenssit ovat aidosti konvekseja, valinnat noudattavat SARP:ia. Jos valinnat siis noudattavat SARP:ia, hyötyteoria on yhtenäinen käyttäytymisen kanssa.

Paljastettujen preferenssien yleistetty aksiooma GARP (General Axiom of Revealed Preference): Jos allokatio X on paljastettu preferoiduksi allokatioon Y ,

ei Y ole koskaan suoraan paljastettu preferoiduksi allokaatioon X. X ei siis ole koskaan valittavissa, kun Y valitaan.

GARP on täysin yleistetty sekä välttämätön ja riittävä ehto hyötyfunktion olemassaololle.

Budjetti	Polettien määrä	Pistemäärä jokaisesta saadusta poletista	
		Diktaattori	Toinen
1	40	3	1
2	40	1	3
3	60	2	1
4	60	1	2
5	75	2	1
6	75	1	2
7	60	1	1
8	100	1	1

Taulukko 1. Eri budjetit ja polettien sekä pisteiden määrä niissä. (Andreoni ja Miller, 1998)

Andreoni ja Miller (1998) suorittivat useita diktaattoripeli -kokeita, joissa toimijat saivat jakaa saamansa ”poletit” itsensä ja toisen pelaajan kanssa. Koehenkilöitä oli yhteensä 142, ja he olivat vapaaehtoisia ylemmän tason taloustieteen opiskelijoita. Polettien määrä oli 100, 75, 60 tai 40 polettia ja poletit olivat joko yhden, kahden tai kolmen pisteen arvoisia. Yksi piste oli arvoltaan 0,10\$. Muuntelemalla polettien määrää ja henkilökohtaista pistearvoa Andreoni ja Miller kykenivät luomaan useita erilaisia tuottoja. Taulukko 1 kuvaa kahdeksan erilaista valintatilannetta, jotka esitettiin koehenkilöille.

Diktaattori tietää jakopäätöksen tehdessään, kuinka monta pistettä kumpikin poleteista saa. Allokatio-ongelma on siis valinta kahden henkilön lopullisten tulojen välillä. Andreoni ja Miller tarkastelivat, oliko näiden kokeiden tuloksilla epäjohdonmukaisuuksia paljastettujen preferenssien vaatimusten kanssa. Testeissä alle 10%:lla oli vähintään yksi rike (13 henkilöllä). Suurimmalla osalla nämäkin olivat kuitenkin hyvin pieniä. Vain yhden koehenkilön valinnat rikkoivat vakavasti GARP -aksiomaa. Andreonin ja Millerin mukaan tämä viittaa siihen, että lähes kaikki koehenkilöt osoittivat rationaalista kysyntää altruismille.

Hyötyfunktio	Sopivuus		Yhteensä
	Vahva	Heikko	
Itsekäs	40	43	83 (47,2%)
Leontief	25	28,5*	53,5 (30,4%)
Täydellinen substituuksi	11	28,5*	39,5 (22,4%)

Taulukko 2. Henkilöiden luokittelu hyötyfunktioiden prototyyppien mukaisesti.

*Yksi henkilö osoitti yhtäläisesti Leontief- ja täydellisen substituuksiin käytöksiä.

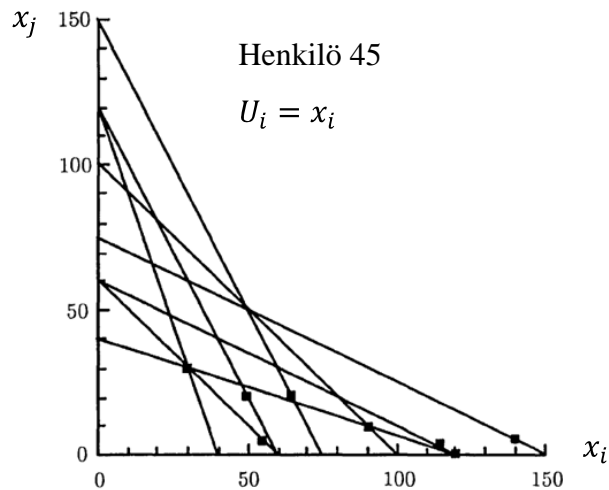
(Andreoni ja Miller, 2002)

Koska lähes kaikkien preferenssit voidaan nähdä rationaalisina, löytyy myös niitä vastaavat hyötyfunktiot. Andreoni ja Miller jakavat koehenkilöiden hyötyfunktiot kolmeen pääryhmään (taulukko 2). Kokeissa 22% henkilöistä oli täysin itsekkäitä pitäen kaikki poletit. Heidän hyötyfunktioitaan kuvaa $U_i=x_i$. Henkilöistä 16% puolestaan jakoi monetaarisen lopputuloksen tasan. Heidän päätöksentekoaan kuvaa erinomaisesti esimerkiksi Leontief -hyöty $U_i=\min\{x_i,x_j\}$. Kolmas ääriyhmä (6%) jakoi poletit maksimoidakseen henkilöiden kokonaistuloksen. Näitä sosiaalisia maksimoijia voidaan kuvata hyötyfunktioilla $U_i=x_i+x_j$.

Loput henkilöt olivat jotain näiden ääripäiden väliltä. Andreoni ja Miller (2002) pyrkivät löytämään myös näille muille sopivia hyötyfunktioita. He kuitenkin tulivat tulokseen, että jo aiemmin nimetyt funktiot kuvasivat myös muiden käyttäytymistä riittävän hyvällä tasolla stereotyyppien luomiseen.

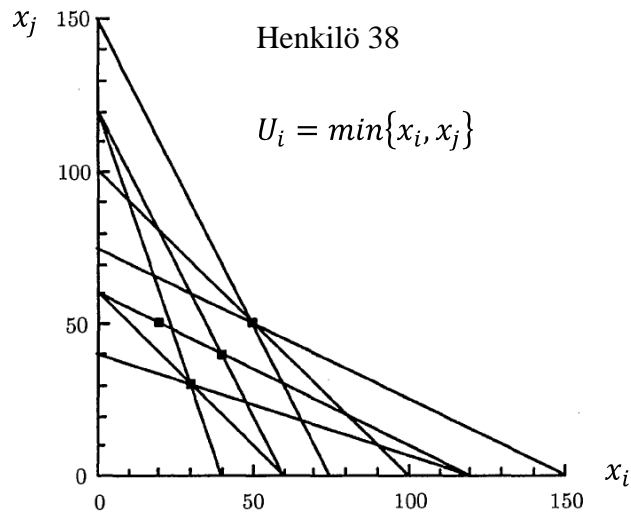
Nämä tulokset ovat tarkastelumme kannalta erityisen mielenkiintoisia monestakin syystä. Tämä koe on edellisen kappaleen empiiristen kokeiden ohella jälleen yksi koe, joka todistaa altruististen insentiivien olemassaolon. Tämän lisäksi se tukee myös kappaleen 3.2 tuloksia egoistien ja altruistien stereotyypeistä. Noin puolet henkilöistä osoittivat selviä altruistisia insentiivejä, kun taas toinen puoli oli puhtaasti itsekästä. Tulevien tarkastelujemme kannalta kuitenkin merkittävintä on, että altruistinen toiminta täyttää hyvinvointifunktion piirteet. Lisäksi mielenkiintoista on, että jopa lähes puolet yksilöistä olivat mallinnettavissa lähes täydellisesti selkeillä ja erittäin yksinkertaisilla hyötyfunktioilla.

Tarkastellaan vielä esimerkkejä eri hyötyfunktioiden mukaan käyttäytyvistä toimijoista ja heidän preferensseistään. Kuva 5 esittää ensin esimerkin heikosti itsekkästä toimijasta. Jokainen budjettisuora valintoja esittävässä kuvassa kuvaa eri budjettien (kts. taulukko 1) suhteellista pistemäärää pelaajien välillä. Mitä suurempi on poletin arvo pelaajalle j suhteessa diktaattoria pelaavaan pelaajaan i , sitä jyrkempi on kulmakerroin ja päinvastoin. Jokaiselle budjettisuoralle on piirretty piste, joka kuvaa henkilön i valintaa kyseisellä budjetilla.



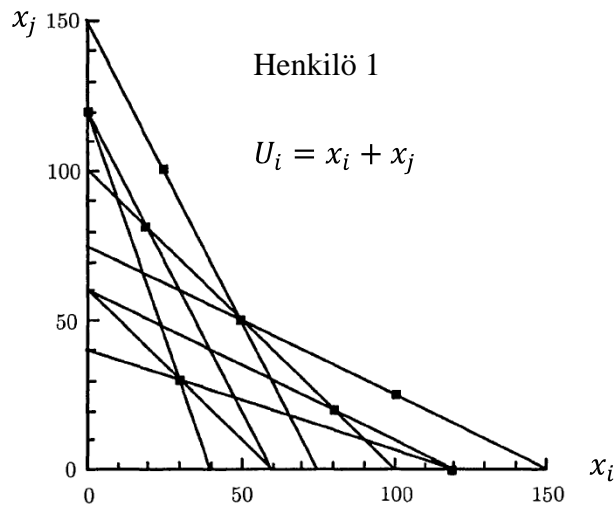
Kuva 5. Heikosti itsekkäitä insentiivejä edustava henkilö diktaattoripelissä. (Andreoni ja Miller, 1998)

Itsekkäällä toimijalla valinnat kuvassa 5 ovat pääsääntöisesti x_i akselin tuntumassa. Vaikka budjettisuoran kulmakerroin muuttuu jyrkemmäksi, ei henkilö 45 allokoij poletteja merkittävästi enempää toiselle pelaajalle. Jonkinasteista hintaherkkyyttä on kuitenkin havaittavissa. Lähes kaikissa valinnoissa henkilö 45 lisäksi osoittaa heikkoja altruistisia insentiivejä allokoijden toiselle pelaajalle ylipäänsä jotain – joskin vain hyvin pieniä summia. Tästä syystä henkilö 45 on vain heikosti itsekäs. Vahvasti itsekkäällä yksilöllä valinnat olisivat aina puhtaasti akselilla x_i .



Kuva 6. Leontief -hyötyfunktioita edustava henkilö diktaattoripelissä. (Andreoni ja Miller, 1998)

Kuvassa 6 on puolestaan esimerkki heikosti Leontief –funktion mukaan toimivasta yksilöstä. Henkilö 38 on lähes jokaisessa tilanteessa valinnut allokaation diagonaalilta eli jakanut pelin monetaarisen lopputuloksen täysin tasan. Yksi tästä eroava ja kuvassakin näkyvä valinta rikkoo WARP –aksioomaa, koska diagonaalilla oleva piste oli valittavissa, kun tästä eroava valinta tehtiin. Tämä kuitenkin Andreonin ja Millerin (1998) mukaan on tulkittavissa pelaajan rehelliseksi virheeksi.

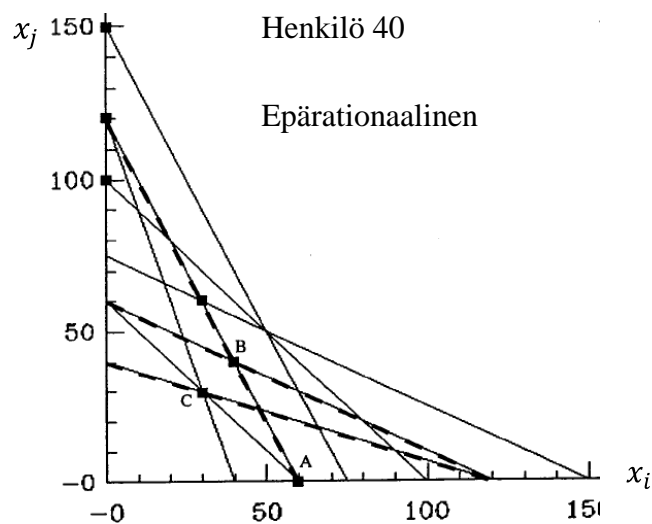


Kuva 7. Sosiaalista maksimoijaa edustava henkilö diktaattoripelissä. (Andreoni ja Miller, 1998)

Kuva 7 on esitys heikosta sosiaalisesti maksimoijasta. Tämä henkilö pyrkii aina painottamaan valintansa siten, että yhteinen monetaarinen tulos olisi mahdollisimman suuri. Heikon toimijasta tekee se, että vaikka kulmakerroin poikkeaisi miinus yhdestä, pelaaja ei pääsääntöisesti allokoisi *ainoastaan* toiselle pelaajalle.

Viimeisenä esimerkkinä on kuvassa 8 Anreonin ja Millerin kokeen ainoa vakavasti GARP -aksioomaa rikkonut *epärationaalinen* toimija. Koehenkilön 40 epärationaalisuus näkyy valinnoissa A, B ja C. Valinta C on tehty toimittaessa loivimmalla, vain pisteen C läpi kulkevalla, budjettisuoralla (olkoon tämä *budjettisuora c*). Valinta A on puolestaan tehty toimittaessa budjettisuoralla, joka kulkee pisteiden A ja B kautta (*budjettisuora a-b*). Valinta B taas on tehty toisella loivemmalla vain B:n kautta kulkevalla budjettisuoralla (*budjettisuora b*). A on suoraan paljastettu preferoiduksi C:hen, koska se on valittu budjettisuoralla a-b ja tästä rajautuva budjettijoukko sisältää pisteen C. C puolestaan on paljastettu

preferoiduksi valintaan A, sillä A on sisällynyt budjettisuoran c rajaamaan budjettijoukkoon. Pelaaja olisi valinnan C tehdessään voinut esimerkiksi myös valita pisteen, jossa hän olisi saanut itse saman monetaarisen tuloksen silti myös pelaajaa j hyödyttäen. Toisaalta pelaaja olisi myös voinut valita tuloksen, joka olisi ollut hänelle itselleen monetaarisesti parempi. Nämä valinnat siis rikkovat WARP -aksioomaa. Lisäksi C on epäsuorasti paljastettu preferoiduksi valintaan B, koska C on valittu myös valinnan A ollessa mahdollinen ja valinta A on tehty, vaikka myös valinta B on ollut mahdollinen. Toisaalta valinta B on suoraan paljastettu preferoiduksi valintaan C valinnan C ollessa budjettisuoran b rajaaman budjettijoukon sisällä. Tämä rikkoo GARP -aksioomaa.



Kuva 8. GARP -rikkeen tehneen ja siten epärationaalisesti käyttäytyneen yksilön valinnat diktaattoripelissä. (Andreoni ja Miller, 1998)

Andreoni ja Miller (1998 ja 2002) näin osoittavat, ettei kaikki epäinstrumentalistinen toiminta ole taloustieteellisen mallintamisen ulkopuolella. He myös painottavat, että tämä mallinnus puoltaa itsekästä Homo Economicus -käsitystä. Epäitsekkäs toiminta

kokeissa voidaan selittää itsekkäiden toimijoiden epäinstrumentalistisen hyödyn maksimointina, kun hyötyfunktiossa nyt kuvautuvat myös altruistiset insentiivit. He siis tunnustavat, etteivät itsekkäät toimijat ole aina instrumentalistisia hyötyjä maksimoivia. Perinteiset rationaalisuuden ja hyödyn käsitteet näyttäytyvätkin nyt Andreonin ja Millerin johtopäätösten pohjalta liian kapeina. Tämä viittaa Kantin (1795) ja Smithin (1759) – sekä monien muiden etenkin perinteisen velvollisuusetiikan kannattajien – osittain jopa naiiviin ja väärään käsitykseen altruismin luonteesta.

4.2. Adam Smithin ongelma

Adam Smith muistetaan taloustieteissä ajatuksistaan taloudentoimijan egoismista ja tämän toiminnan yhteiskunnallisesta tehokkuudesta (*The Wealth of Nations*, 1776). Ajatuksia egoismista toki ovat esittäneet muutkin suuret ajattelijat. Esimerkiksi Nietzsche toteaa teoksessaan *Antikristus* (1908), että ”rappeutunein ihmislaji on pappi: hän opettaa luonnonvastaisuutta...”, sillä, kuten Nietzsche esittää, ”kristinusko on säälin uskonto”. Tämä näkemys kristinuskosta pohjautuu Nietschen halveksimien ominaisuuksien kuten huono-osaisten auttamisen ja empatian korostamiseen kristinuskossa. Ovathan nämä kaikkea muuta kuin rationaalista oman edun tavoittelua. Toisin kuin Nietzsche, Adam Smith kuitenkin painotti egoismissakin yhteisen hyvän edesauttamista. Lisäksi hän ennen *The Wealth of Nations* -kirjan kirjoittamista oli jo kirjassaan *Theory of Moral Sentiments* (1759) kuvannut universaalien altruistien, joka empatiantuntonsa kautta edesauttaa myös yhteistä hyvää. Näiden kahden Adam Smithin esittämän ihmiskuvan näennäisen ristiriitaisuuden takia taloustiede lieneekin omaksunut vain toisen Smithin näkemyksistä – rationaalisen

egoistin. Tämä näkemys mahdollisti luontevasti markkinoita ohjaavan ”näkyttömän käden” hyväksymisen ja ihminen toimijana kadotettiin (Kyrö, 2006).

Tässä kappaleessa pyrin tarkastelemaan Smithin kahden ihmiskäsityksen ristiriitaa sekä Kolmin (2006) esittämää ratkaisua. *Theory of Moral Sentiments* -kirjan altruisti on kuitenkin moraalinen vain rajatussa määrin. Suuri osa hänen altruismistaan kumpuaa hedonismista, kuten tässäkin työssä on jo useaan otteeseen ehdotettu (esimerkiksi neurotalous-tieteen tuloksia tarkasteltaessa). Hän arvostaa toisen ihmisen mielihyvää, koska empatian osoittaminen ja toisaalta hyvinolontunteen tarttuminen (contagion) palvelevat hänen omaa mielihyväänsä (Kolm, 2006). Ylipäänsä tällaisen käytöksen kutsuminen moraaliseksi on ongelmallista – etenkin tuon ajan velvollisuusetiikan näkökulmasta, jonka edustajaksi esimerkiksi Smithin aikalainen Kant jo aiemmin todettiin (1795). Smithin altruisti ei siis empatian tunnostaan ja Smithin moraaliseksi kuvaamasta luonteestaan huolimatta ole välttämättä täysin egoistisen toiminnan ulkopuolella. Hän ruokkii omaa mielihyväänsä altruistisin toimin.

Kolmin (2006) mukaan Smithin kaksi ihmiskäsitystä eivät muodosta ristiriitaa kolmesta syystä:

1. Vaikka markkinat nojaavat itsekkääseen toimintaan, altruismi on osittain välttämätöntä niiden toiminnalle. Ihmisen altruistisen luonteen takia markkinoiden toimintaa järkyttävä varastaminen ja huijaaminen ei ole yleistä.
2. Altruismi saattaisi vaarantaa hinnan määräytymisen tehokkuuden kilpailun puutteen, hinnan alennusten ja ylisuurien maksujen takia. Suurin osa perheiden

ulkopuolisesta altruismista kuitenkin ohjautuu yhteiseen tulonjakoon köyhiä kohtaan. Tämä on helpointa ja tehokkainta suorittaa julkisen sektorin toimesta, jolloin markkinoilla ihmiset voivat toteuttaa itsekkäitä tarkoituksiaan huoletta.

3. Kaupanteko, antaminen ja tulonjaosta äänestäminen tapahtuvat eri tilanteissa, eri aikoina ja eri sidosryhmien kanssa – markkinat, perhe, hyväntekeväisyys, valtiomuoto.

Enempää Kolmin (2006) ratkaisupyrkimystä tarkastelematta voitaneen huomata kohtien 1-3 perusteella altruismin itseasiassa tukevan itsekkyyden tehokkaaksi tekemiä markkinoita. Tämän se tekee julkisen sektorin auttamana suoraan koko yhteisön hyväksi. Tällöin se ei järkytä vapaita markkinoita vaan pikemminkin puhdistaa ihmisten omantunnon itsekkyyden harjoittamiseksi tuossa ympäristössä.

Kappaleiden 3.2.2 ja 3.2.4 teorioiden altruistit ja egoistit stereotyyppinä voitaneen myös nähdä mahdollisina selittäjinä Adam Smithin ongelmaan ja tukijoina Kolmin ehdottamaan ratkaisuun. Mikäli pelien säännöt ovat erilaiset esimerkiksi yhteiskuntajärjestyksen ylläpidossa ja markkinatalouden luomassa kuluttaja- tai työympäristössä, on hyvin mahdollista ja perusteltua, että niissä toimimisen normit muodostuvat erilaisiksi toisen painottaessa yhteistyötä ja toisen oman edun tavoittelua.

Tämä ratkaisu osoittaa, ettei valintaa ihmisen käyttäytymisen selittäjästä tarvitse tehdä. Rationaalisuus voi selittää altruistisen käyttäytymisen, mutta todellisen altruismin olemassaolo ihmisluonteen piirteenä ei ole ongelma taloustieteelle – vain haaste. Taloustieteen ei siis tarvitse olla ns. ”pimeä tiede” (dismal science)

inhorealisticine ihmiskäsityksineen, kuten Nietzschen (1908) ylevä ihminen. Sitaattia Kolmilta (2006) lainatakseni ”man is neither angel nor beast”.

Näin siis kappaleessa 4 on vastattu kysymyksenasettelun toiseen kysymykseeni niin teoreettisella kuin semanttisellakin tasolla. *Altruistinen käyttäytyminen voidaan nähdä rationaalisena toimintana.*

5. Teorioita altruistisista hyötyfunktioista

Empiiriset tulokset altruismin olemassaolosta ja mahdollisuudesta kuvata se osana hyötyfunktioita ovat johtaneet useisiin erilaisiin yrityksiin luoda rationaalinen hyödyn maksimoinnin teoria, joka voisi selittää saatuja tuloksia. Näitä teorioita tarkastellessani kiinnitin huomioni malleihin, joita Fehr ja Schmidt (2006) kutsuvat sosiaalisten preferenssien (social preferences) malleiksi. Tarkastelun alaisina malleina olivat Andreonin ja Millerin (1998) GARP koe sekä Fehr ja Schmidtin (1999) sekä Boltonin ja Ockenfelsin (2000) epätasa-arvon välttämiseen perustuva esitys (inequity aversion). Myös Fehr ja Schmidt (2006) tarkastelevat artikkelissaan edellä mainittuja teorioita mutta erottavat myös kaksi muuta pääsuuntausta: toisistaan riippuvaiset preferenssit (interdependent preferences) ja aikomuksille perustuvan vastavuoroisuuden (intention based reciprocity). Pyrin tässä kappaleessa esittelemään edellä mainittuja hyötyfunktioita pinnallisella tasolla sekä Fehr ja Schmidtin näkemyksiä näiden teorioiden hyvyydestä.

5.1. Sosiaalisten preferenssien malli

Sosiaalisten preferenssien mallit lähtevät liikkeelle ajatuksesta, että päätöksentekijä saattaa välittää paitsi omasta taloudellisesta hyödystään, myös muille allokoituista materiaalisista resursseista. Fehr ja Schmidt (2006) huomauttavat omassa määritelmässään sosiaalisista preferensseistä, että rationaalisen toiminnan käsite on perinteisessä itsekkäässä allokaatiossa sosiaalista preferenssiä kapeampi käsite. Myös Fehr ja Schmidt viittaavat siis laajempien rationaalisuuden ja hyödyn käsitteiden tarpeeseen altruistisen hyötyfunktion muodostamiseksi.

Kappaleessa neljä esitetty altruismin rationaalisuustarkastelu (Andreoni ja Miller, 1998 ja 2002) oletti altruismin olevan nimenomaan sosiaalisten preferenssien maksimointia. Henkilö on siis altruistinen, jos $\partial U/\partial x_j$ (missä j on toinen yksilö) on aidosti positiivinen eli yksilön hyöty kasvaa muiden hyödyn kasvaessa. Tosin Andreonin ja Millerin tarkastelussa suurin osa koehenkilöistä osoittautui melko itsekkäiksi. Altruistisetkin koehenkilöt kuitenkin toimivat rationaalisesti kuluttajanteorian valinnan aksioomien mukaan. Koska tämä tarkastelu todistaa hyötyfunktion olemassaolon, se antaa myös perustan hyötyfunktion etsimiselle.

Formaalimman esityksen mahdollistamiseksi merkitään yksilöitä joukolla $\{1,2,\dots,N\}$ ja resurssien allokaatiota joukosta X vektorilla $x=(x_1,x_2,\dots,x_N)$. Kuten aiemmin, merkitään, että x_i on henkilön i monetaarinen tulos. Itsekkään toimijan hyöty riippuu siis termistä x_i . Nyt henkilöllä i voidaan sanoa olevan sosiaalisia preferenssejä, jos henkilön i hyötyyn vaikuttaa muutos toisen henkilön tuloissa eli termissä x_j , jossa siis $j \neq i$.

Fehr ja Schmidt (1999) esittävät oman hyötyfunktionsa, joka voisi tuottaa kokeissa havainnoituja tuloksia. Heidän mallinsa perustuu oletukselle, että hyöty on joko monotonisesti kasvava tai monotonisesti vähenevä muiden yksilöiden hyödyn suhteen. Yksilö on Fehrin ja Schmidtin oletuksen mukaan altruistinen muita yksilöitä kohtaan, jos heidän materialistiset tulonsa ovat jonkin kohtuullisen kiintopisteen alapuolella. Jos toisten tulot taas ovat tämän kiintopisteen yläpuolella, tuntee yksilö kateutta. Fehrin ja Schmidtin näkemys pohjautuu heidän käsitykseensä altruistisen rangaistuksen olemassaolosta (Fehr ja Gächter, 2002). Useimmille koeasetelmille on turvallista olettaa kohtuullisen allokaation vastaavan monetaarisen lopputuloksen

jakamista tasan pelaajien kesken. Näiden oletuksien pohjalta Fehr ja Schmidt esittävät hyötyfunktion

$$U_i(x) = x_i - \alpha_i \frac{1}{N-1} \sum_{j \neq i} \max\{x_j - x_i, 0\} - \beta_i \frac{1}{N-1} \sum_{j \neq i} \max\{x_i - x_j, 0\},$$

joka pyrkii yksinkertaistamaan heidän ideansa. Hyötyfunktiossa $0 \leq \beta_i \leq \alpha_i$, ja $\beta_i \leq 1$. Henkilö on nyt altruistinen eli $\partial U / \partial x_j \geq 0$ jos ja vain jos $x_i \geq x_j$. Mikäli henkilön hyvinvointi on huonompi kuin toisella osapuolella, ei henkilö ole altruistinen. Olennaista mallissa on myös, että hyötytappio epätasa-arvosta on suurempi, jos toisen henkilön tulo on suurempi kuin pelaajan i , kuin, jos toisen henkilön tulo on pienempi ($\alpha_i \geq \beta_i$). Yksinkertaistettuna kahden pelaajan tapaukseen, jollaisia tarkastelemamme kokeet ovat olleet, Fehr ja Schmidt (1999) muokkaavat hyötyfunktioitaan. Tällöin

$$U_i(x) = x_i - \alpha_i \max\{x_j - x_i, 0\} - \beta_i \max\{x_i - x_j, 0\}, i \neq j.$$

Tämä hyötyfunktio pystyy rationalisoimaan positiivisen altruistisen toiminnan kuten diktaattoripelien avokätisyyden. Toisaalta se pystyy selittämään myös altruistiset rangaistukset kuten uhkavaatimuspelien pienten tarjousten torjumisen. (Fehr ja Schmidt, 2006) Fehr ja Schmidt (1999) myös osoittavat, että uhkavaatimuspelien tuloksista saatujen α :n ja β :n jakaumat voivat yllättävän tarkasti ennustaa monien muidenkin pelien tuloksia.

Bolton ja Ockenfels (2000) esittävät samantapaisen epätasa-arvon välttämisen hyötyfunktion, jonka he niinkään osoittavat kykenevän selittämään useita epäitsekkäitä pelituloksia.

Jokainen pelaaja i pyrkii heidän mukaansa maksimoimaan *motivaatiofunktioitaan*

$$U_i = U_i(x_i, \sigma_i), \text{ missä}$$

$$\sigma_i = \sigma_i(x_i, c, N) = \begin{cases} \frac{x_i}{c}, & \text{jos } c > 0 \\ \frac{1}{N}, & \text{jos } c = 0 \end{cases}$$

on pelaajan i suhteellinen osuus tulosta ja

$$c = \sum_{j=1}^N x_j$$

on rahallinen kokonaistulo.

Annetulle σ_i hyötyfunktio oletetaan x_i :n suhteen kasvavaksi ja konkaaviksi. (Fehr ja Schmidt, 2006) Tämä tarkoittaa, että jokaiselle annetulle suhteelliselle tulolle, pelaajan i :n valinnat ovat johdonmukaisia normaalien rahalle tehtyjen preferenssioletusten suhteen. Annetulle x_i hyötyfunktio puolestaan on aidosti konkaavi pelaajan i kokonaistulon σ_i suhteen. Lisäksi oletetaan että hyöty saavuttaa maksiminsa, kun $\sigma_i = 1/N$, kun muille ei ole tarjolla hyötyjä ($c = 0$), muutoin se on suhteessa muiden hyötyihin (Fehr ja Schmidt, 2006).

Esitettyjen oletusten perusteella on siis olemassa maksimipiste, jota Bolton ja Ockenfels (2000) kutsuvat sosiaalisiksi referenssipisteeksi. Tässä pisteessä yksilön osuus tulosta on sama kuin keskimääräinen osuus. Ennen tätä pistettä motivaatiofunktio U_i on termin σ_i suhteen kasvava ja yksilö on kateellinen. Sosiaalisen referenssipisteen jälkeen yksilö puolestaan on altruistinen ja motivaatiofunktio U_i on termin σ_i suhteen vähenevä.

Erona Fehrin ja Schmidtin malliin on suhde muihin pelaajiin. Fehrin ja Schmidtin painottaessa yksilön i yksilöllistä suhtautumista jokaiseen yksittäiseen kohteeseen Bolton ja Ockenfels näkevät yksilön kiinnostuksen kohdistuvan pikemminkin muihin yleensä, jolloin tarkastelu kohdistuu keskivertoyksilöön.

Derivaattojen $\partial U / \partial \sigma_i$ ja $\partial U / \partial x_j$ positiivisuus ja negatiivisuus riippuu siis siitä saako henkilö enemmän vai vähemmän kuin keskiarvo. Boltonin ja Ockenfelsin mallissa hyöty ei siis riipu yksilön j suhteesta yksilöön i, vaan pikemminkin yksilön i suhteesta keskivertotasoon. (Fehr ja Schmidt, 2006)

Bolton ja Ockenfels (2000) käyttävät myös mielenkiintoisesti termiä *motivaatiofunktio* hyvinvointifunktion sijaan. Tämä onkin erittäin mielenkiintoinen valinta. Motivaatiofunktiolla Bolton ja Ockenfels taitavasti välttävät ristiriidan altruismin ja hyödyn käsitteiden välillä.

Charness ja Rabin (2002) esittävät myös oman ratkaisunsa hyötyfunktioksi yhdistämällä yhteiskunnan hyvinvointifunktion yksilön hyvinvointifunktioon maksimoitavaksi tekijäksi. Tällöin heidän hyötyfunktionsa

$$U_i(x_1, x_2, \dots, x_N) = (1 - \gamma) x_i + \gamma W(x_1, x_2, \dots, x_N), \text{ missä } \partial W / \partial x_j \geq 0.$$

Sosiaalisten preferenssien mallit yhdistävät yksilön hyötyfunktioon yksilön oman materiaalisen hyvinvoinnin lisäksi myös muiden yksilöiden materiaalisen hyödyn (Fehr ja Schmidt, 2006). Erona aikaisempiin malleihin tällöin on, että yksilön hyöty on aina kasvava muiden tulojen suhteen ja yksilö on siis aina altruistinen.

5.2. Toisistaan riippuvaiset preferenssit

Toisistaan riippuvaisten preferenssien mallit (models of interdependent preferences) olettavat yksilöiden olevan kiinnostuneita toisten henkilöiden ”tyypistä” (Fehr ja Schmidt, 2006). Malleissa oletetaan henkilön olevan aina joko altruistinen tai itseks. Jos altruistinen henkilö tietää olevansa kanssakäymisessä toisen altruistisen toimijan kanssa, hänen oma toimintansa on myös altruistista ja pelien tapauksessa avokätistä. Jos altruistinen henkilö puolestaan tietää olevansa tekemisissä itsekkään toimijan kanssa, muuttuvat myös hänen preferenssinsä itsekkäiksi. Pelaajan i preferenssit riippuvat siis siitä, ovatko pelaajan j preferenssit itsekkäät vai altruistiset.

Nämä toisistaan riippuvaisten preferenssien mallien oletukset herättävät osin epäilyksiä. Yksilöiden oletetaan olevan ”absoluuttisesti” joko altruistisia tai itsekkäitä. Toiseksi mallissa oletetaan, että henkilöt tietävät toisten todelliset preferenssit. Seuraavaksi esiteltävä aikeisiin perustuva vastavuoroisuus on osin samankaltainen toisistaan riippuvaisten preferenssien kanssa, mutta tässä mallissa on painotettu etenkin yksilön tulkintaa – ei tietoa – toisen aikeista.

5.3. Aikeisiin perustuva vastavuoroisuus

Aikeisiin perustuvan vastavuoroisuuden (intention based reciprocity) mallit olettavat yksilön välittävän nimenomaan toisten aikeista (Fehr ja Schmidt, 2006). Jos yksilö kokee toisen haluavan kohdella häntä hyvin, toimii hän vastavuoroisen altruistisesti toista kohtaan. Vihamielisinä koettuihin aikeisiin puolestaan vastataan myös vihamielisesti. Kyse on siis todella kokemuksista eikä yksilöiden todellisista luonteista. Fehr ja Schmidt esittävät näiden mallien vaativan psykologista taustaa, sillä perinteinen peliteoria ei kykene mallintamaan syntyvää kahden mahdollisen

tasapainon tilaa. Kaksi altruistista yksilöä voi päätyä yhtä hyvin vihamieliseen ratkaisuun kuin altruistiseenkin – kyse on henkilöiden kyvyistä tulkita toisen aikeita.

Muun muassa Charness ja Rabin (2002) ovat kehittäneet malliaan yhdistämällä siihen aikeet. Fehr ja Schmidt (2006) kuitenkin moittivat mallia sen monimutkaisuudesta ja siten implementoinnin vaikeudesta.

Tämän kappaleen malleista etenkin sosiaalisten preferenssien mallit ovat osoittaneet kykynsä rationalisoida altruistisen käyttäytymisen aiheuttamat kokeelliset tulokset. (Fehr ja Schmidt, 2006) Kaksi jälkimmäistä tapaa mallintaa altruistisia preferenssejä ovat kuitenkin kompleksisuudessaan mahdollisesti parempia tuottamaan tarkempia tuloksia, jos niiden ongelmat pystytään ratkaisemaan. Vaikka malleihin ei tarkalla matemaattisella tasolla mentykään, olemme nähdäkseni todenneet kysymyksenasettelun kolmannen ja viimeisen väitteeni. *Altruismi on mahdollista huomioida käytännön tasolla taloustieteen mallinuksissa, sovelluksissa ja hyötyfunktioissa.*

6. Johtopäätöksiä

Altruistisen käyttäytymisen olemassaolosta empiiristen todisteiden valossa ei voi enää taloustieteen kentässä väitellä, kuten kappaleessa kolme osoitettiin. Kuitenkin sen perimmäisestä luonteesta edelleen kiistellään. Taloustieteen perinteisen näkemyksen mukaan ihmisen perimmäinen luonne on itsekäs ja rationaalisesti materialistisia hyötyjään maksimoiva. Kappaleessa neljä tarkasteltiin Andreonin ja Millerin näkemyksiä altruismin rationaalisesta ja itsekkästä olemuksesta. Heidän oletuksistaan ja lopputuloksista nähtiin, että rationaalisuus ja hyöty näyttävät tämänkaltaisessa tarkastelussa ihmisen tunteet ja moraalin huomioivina käsitteinä toisin kuin perinteisesti on ehdotettu. Myös kappaleen viisi hyötyfunktioiden tarkastelut ottivat lähtökohdaksi nimenomaan yksilön suoran preferenssin altruismiin – ei sen materialistisiin vaikutuksiin.

Saatuja tuloksia tukevat neurotaloustieteen tulokset, jotka osoittavat altruistisen käyttäytymisen aiheuttavan yksilölle suoria hedonistisia kokemuksia. Näiden tulosten valossa on perusteltua nähdä yksilö oman hyötynsä rationaalisena maksimoijana myös käyttäytyessään altruistisesti. Kuitenkin erona perinteiseen itsekkääseen näkemykseen tällöin on yksilön luontainen halu todelliseen altruismiin ilman materialistisia pyyteitä tai odotuksia.

Altruistiset insentiivit eivät kuitenkaan jää vain satunnaisiksi yksilöllisiksi preferensseiksi, vaan kuten kappaleessa 3.2.4 todettiin, ne ovat keskeinen osa koko yhteisön toimintaa ja jatkuvuutta. Klassisen taloustieteen puolustajat ovat usein viimeisenä oljenkortenaan tarttuneet perusteeseen, että taloustieteiden tehtävä ei ole katsoa yksilöitä, vaan pikemminkin taloudentoimijoiden agregaatteja. Gintisin

evolutiivisen mallin valossa tämäkin peruste näyttää hataralta. Mikäli altruismi on välttämätön ja merkittävä osa yhteisöjä, se ei katoa toimijoiden hyötyfunktioita agregoidessa vaan nimenomaan tulee osaksi sitä.

Nämä huomiot eivät kuitenkaan tee taloustieteen tekemisestä mahdotonta tai sen keskeisimmistä teorioista vääriä. Ne pikemminkin asettavat sille vain uusia haasteita, jotka on mahdollista mallintaa osana taloustieteen teorioita ja huomioida niitä tulkittaessa. Tärkeänä aihetta pidän sen takia, että kritiikinalainen Homo Economicus -oletus on koko taloustieteen tieteenteorian keskiössä ja ensimmäinen tukijalka koko tieteenhaaralle. Näin ollen sen epämääräinen määrittely tai väärin tulkitseminen heijastuu kaikkiin siitä johdettuihin taloustieteen teorioihin.

Alussa esittämäni kolmen kohdan tavoite täyttyi. Tämä tarkoittaa, että altruismin huomioiminen yksilön toiminnan taloudellisessa tarkastelussa on paitsi *tarpeellista* myös *mahdollista* sekä *hyödyllistä*. Näin ollen uskon taloustieteen käyttämän termistön ja perusaksiooman olevan uuden, täsmällisemmän määrittelyn tarpeessa.

Ehdotukseni Homo Economicus -oletuksen ympärille rakentuvasta semantiikasta on lyhyesti seuraava:

Homo Economicus näyttäytyy tarkastelujemmekin jälkeen edelleen hyötyään rationaalisesti maksimoivana yksilönä. Tässä määritelmässä ei kuitenkaan ole sijaa väitteelle *egoismista*, sillä jopa puolet yksilöistä osoittaa yleensä luontaista kysyntää altruismille. *Hyöty* puolestaan on nyt laajennettava pelkkien instrumentalististen hyötyjen ulkopuolelle. Myös moraalit voidaan laskea yksilön hyötyä nostavaksi. *Altruismin* kantilaisesta määritelmästä on myös tingittävä. Altruismia ei tarvitse

tulkita kaikesta mielihyvystä kieltäytyvänä toimintana, vaan pikemminkin terminä taloudellisessa mielessä pyyteettömälle toiminnalle. Ollessaan osa hyötyfunktioita, se kuitenkin on yksilön hyötyyn ja onnellisuuteen vaikuttava tekijä. Mistä intentio noudattaa omaa moraalialuekumpuaa, ei kuitenkaan ole mielekäs tai oleellinen kysymys taloustieteelle.

Rationaalisuus sen sijaan ei käsitteenä välttämättä vaadi juurikaan uudelleenmäärittelyä, mikäli voimme hyväksyä aiemmin esittämäni ”teknisen” määritelmän. Rationaalisuus korkeimman hyödyn valintana on selkeä eikä se pidä sisällään arvovalintoja. Sen sijaan egoismin ja rationaalisuuden käsitteiden sotkeminen ei ole oikeutettua, eikä tällaisia virheitä tulisi etenkin tieteellisissä teksteissä missään nimessä hyväksyä.

Siinä, missä ensimmäinen motivaationi oli lähteä kaatamaan monien tieteiden virheelliseksi katsomaa ihmiskuvaa, olen nyt vakuuttunut sen toimivuudesta. Ihminen todella voidaan edelleen nähdä omaa hyötyään ajavana yksilönä, kunhan huomioidaan myös mahdollisuus epäinstrumentalistisiin hyötyihin. Tässäkin näkemyksessä on omat uhkansa. Taloustiede on eksakti ja pedantti tiede. Se ei saa muuttua filosofiseksi kielileikiksi. Toisaalta juuri siksi siltä pitää voida odottaa ja vaatia hyvin jäsenneilyä ja yhteistä näkemystä myös sen tieteenteoriasta ja semantiikasta.

LÄHTEET

Andreoni, J. ja Miller, J. (1998). Analyzing choice with revealed preference: Is altruism rational? *The Handbook of Experimental Economics Results*. Charles, P ja Vernon, S.

Andreoni, J. ja Miller, J. (2002). Giving according to GARP: An experimental test of the rationality of altruism. *Econometrica* 70, 737-753.

Bolton, G.E. ja Ockenfels, A. (2000). A theory of equity, reciprocity and competition. *American Economic Review* 100, 166-193.

Boyd, R. (1982). Density dependent mortality and the evolution of social behavior. *Animal Behav.* 30, 982.

Brown, M., Falk, A. ja Fehr E. (2004). Relational Contracts and the Nature of Market Interactions. *Econometrica*. *Econometric Society*, 72(3), 747-780.

Charness, G. ja Rabin, M. (2002). Understanding social preferences with simple tests. *Quarterly Journal of Economics* 117, 817-869.

Elster, J. (2006). Altruistic Behavior and Altruistic Motivations. *Handbook of The Economics of Giving, Altruism and Reciprocity, Volume 1 Foundations, Handbooks in Economics* 23. Kolm ja Ythier (2006).

Falk, A. ja Fischbacher, U. (2006). A theory of reciprocity. *Games and Economic Behavior* 54, 293-315.

Fehr, E. ja Falk, A. (1999). Wage Rigidity in a Competitive Incomplete Contract Market. *Journal of Political Economy*, University of Chicago Press, 107(1), 106-134.

Fehr, E. ja Fischbacher, U. (2003). The Nature of Human Altruism. NATURE Vol. 425.

Fehr, E. ja Fischbacher, U. (2004). Third party punishment and social norms. Evolution and human behavior 25, 63-87.

Fehr, E. ja Gintis, H. (2007). Human Motivation and Social Cooperation. Annual Review of Sociology 33,3 1-22.

Fehr, E. ja Gächter, S. (2000). Cooperation and Punishment in Public Goods Experiments. Paper provided by CESifo Group Munich in its series CESifo Working Paper Series with number CESifo Working Paper No. 183.

Fehr, E. ja Gächter, S. (2002). Altruistic punishment in humans. NATURE Vol. 415.

Fehr, E., Kirchsteiger, G. ja Riedl, A. (1993). Does Fairness Prevent Market Clearing? An Experimental Investigation. The Quarterly Journal of Economics, MIT Press, 108(2), 437-59.

Fehr, E. ja Schmidt K.M. (1999). A theory of fairness, competition and cooperation. Quarterly Journal of Economics 114, 817-868.

Fehr, E. ja Schmidt K.M. (2006). The Economics of Fairness, Reciprocity and Altruism – Experimental Evidence and New Theories. Handbook of The Economics of Giving, Altruism and Reciprocity, Volume 1 Foundations, Handbooks in Economics 23. Kolm, S.C. ja Ythier, J.M. (2006).

- Fischbacher, U., Gächter, S. ja Fehr, E.** (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economic Letters* 71, 397-404.
- Gintis, H.** (2000). Strong Reciprocity and Human Sociality. *Journal of Theoretical Biology* 206, 169-179.
- Kant, I.** (1785 – 1788) Siveysopilliset pääteokset. (Tapojen metafysiikan perustus, *Grundlegung zur Metaphysik der Sitten*, 1785; Käytännöllisen järjen kritiikki, *Kritik der praktischen Vernunft*, 1788.) Suomentanut J. E. Salomaa. (1. painos 1931).
- Kant, I.** (1795). Ikuiseen rauhaan. Näköispainos alkuperäisteoksesta Karisto Oy 4. painos (2000).
- Kolm, S-C.** (2006). Introduction to the Economics of Giving, Altruism and Reciprocity. *Handbook of The Economics of Giving, Altruism and Reciprocity, Volume 1 Foundations, Handbooks in Economics* 23. Kolm ja Ythier (2006).
- Kyrö, P.** (2006). The Transitional Development of Entrepreneurship- Dialogue Between New Economic Activity, Work and Freedom. *Estudios de Economía Aplicada* Vol. 23-1 (2006).
- Nietzsche, F.** (1908). Antikristus. Näköispainos alkuperäisteoksesta. *Unio Mystica* (2001).
- Pihlanto, P.** (2002). Talous: Rationaalisen toimijan myytti taloustieteissä, yrity maailmassa ja yhteiskuntapolitiikassa. *Niin&Näin* 2/2002.

Rabin, M. (1993). Incorporating Fairness into Game Theory and Economics. American Economic Review, 83, 1281-1302.

Smith, A. (1759). The Theory of Moral Sentiments. E-book.

Smith A. (1776). An Inquiry into the Nature and Causes of the Wealth of Nations. E-book.

Solarz, B. (2006). Putting the 'Self' in Selfless. Yale Economic Review summer 2006.

Staveren, I. (2001). The Values of Economics. An Aristotelian Perspective. Routledge.

Vihanto, M. (2005). Rationaalisuusoletuksen valinta oikeustaloustieteessä. <http://www.tukkk.fi/yltalous/kt/mvihanto/ratval.pdf>.

WILSON, D. S. (1977). Structure demes and the evolution of group-advantageous traits. Am. Nat. 111, 185.

Zafirovski, M. (2003). Human Rational Behavior and Economic Rationality. Electronic Journal of Sociology.

LIITE 1.

Insentivi tunne kuten rangaistuksen pelko tai palkkion odotus, joka motivoi toimintaa ja nähtyä vaiivaa.

Instrumentalistinen tavoite kohdistuu konkreettiseen ja mitattavaan välittömästi tai välillisesti materialistiseen hyötyyn.

Intentio ihmisen mentaaliin perustoimintoihin kuuluva tarkoitus tai pyrkimys.

Fundamentalistinen dualismi on vahva idealistinen suhtautuminen todellisuuden kaksijakoisuuteen.

Objektivismi on ideologia, joka painottaa asioiden riippumattomuutta ihmisistä.

Ontologia eli **oppi olevaisesta** on filosofian ja erityisesti niin kutsutun metafysiikan osa, joka tarkastelee todellisuuden yleistä luonnetta ja pohtii, millainen ympäröivä todellisuus on. Ontologian perimmäinen kysymys kuuluu: ”Mitä on olemassa?”

Realismi on aistimaailman objektien todellista olemassaoloa kannattava ontologian suuntaus.

Relativismi suhteellisuusajattelu on nimitys useille käsityksille, joita yhdistää näkemys siitä, että jokin asia on suhteellinen johonkin toiseen seikkaan nähden tai tällaisesta riippuvainen.

Semantiikka eli merkitysoppi tutkii esimerkiksi sanojen ja ajatusten merkitystä.

Subjektivismi on objektivismin vastakohta painottaen asioiden riippuvuutta ihmisestä.

Paradigma tarkoittaa oikeana pidettyä, yleisesti hyväksyttyä ja auktoriteetin asemassa olevaa teoriaa tai viitekehystä, kuten talouden toimijan egoismia, evoluutioteoriaa tai suhteellisuusteoriaa.

Positivismin mukaan oivalletaan vähitellen, että metafysiset kysymykset ovat ratkaisemattomia ja jäävät sellaisiksi. Me voimme saada tietoa vain asioista, joita havaintomme pystyy tavoittamaan. Tieteen tehtäväksi tulee selvittää näiden tosiasioiden kesken vallitseva lainalainen yhteenkuuluvuus ja siten ennustaa tulevia tapahtumia.