

Kuinka tämän selittäisi

Kausaaliselityksen manipulaatioteoria ja sen soveltaminen taloustieteessä

Jonatan Kanervo
Pro gradu
Teoreettinen filosofia
Helsingin yliopisto
Humanistinen tiedekunta
Helmikuu 2021
Ohjaaja: Samuli Reijula

Tiedekunta - Fakultet - Faculty Humanistinen tiedekunta		Koulutusohjelma – Utbildningsprogram – Degree Programme Filosofian maisteriohjelma	
Tekijä - Författare - Author Atte <u>Jonatan Kanervo</u>			
Työn nimi - Arbetets titel Kuinka tämän selittäisi – Kausaaliselityksen manipulaatioteoria ja sen soveltaminen taloustieteessä			
Oppiaine - Läroämne - Subject Teoreettinen filosofia			
Työn laji/ Ohjaaja - Arbetets art/Handledare - Level/Instructor Pro gradu -tutkielma / Samuli Reijula		Aika - Datum - Month and year Helmikuu 2021	Sivumäärä - Sidoantal - Number of pages 84
Tiivistelmä - Referat - Abstract			
<p>Tutkielma tarkastelee James Woodwardin kausaaliselityksen manipulaatioteoriaa ja sen soveltamista taloustieteessä. Se keskittyy luomaan filosofisesti perustellun ja käytäntöä lähellä olevan tulkinnan kausaalisesta selittämisestä taloustieteen alalla.</p> <p>Kausaatiolla ja selityksellä on keskeinen rooli ymmärryksen tuottamisessa. Niillä on lisäksi korvaamatonta käytännöllistä merkitystä, sillä tieto kausaalisuhteista mahdollistaa tapahtumien kontrolloinnin. Tästä syystä suuri osa tieteellisestä toiminnasta keskittyy kausaalisuhteiden löytämiseen, ja niiden laajuuden arviointiin. Tieteellisen työn tarkkuuden takaamiseksi kausaatio, ja sen varaan pohjautuvat selitykset, vaativat tarkan määritelmän. Tutkielma esittelee Woodwardin manipulaatiosuhteisiin perustuvan määritelmän ja selvittää, miten tästä kausaation käsitteestä voidaan johtaa selityksiä. Tähän esitykseen perustuen tutkielma tarkastelee lisäksi Woodwardin teorian käyttökelpoisuutta taloustieteen metodologian tulkitsemisessä. Yhtymäpinnan Woodwardin teorian ja taloustieteen metodologian välille luo Uskali Mäen eristyksiin perustuva näkemys taloustieteellisestä mallintamisesta.</p> <p>Ensimmäinen luku esittelee keskeiset rajaukset ja lähtökohdat tutkimukselle. Toinen luku käsittelee Woodwardin näkemystä kausaatiosta ja antaa tarkan määritelmän kausaatiolle tapahtumien välisten kontrafaktuaalisten riippuvuussuhteiden kautta. Kolmas luku selvittää, miten edellä määritellystä kausaation käsitteestä voidaan johtaa selityksiä, ja perustelee, miksi esiteltävä kausaaliselityksen teoria on ratkaiseviltä osin parempi kuin useat historialliset selitysteoriat. Neljäs luku esittelee Mäen näkemyksen malleista sijaissysteemeinä ja argumentoi, että esiteltävä näkemys taloustieteellisestä mallintamisesta voidaan nähdä taloustieteen tarkoituksiin sopivana käytännön sovelluksena Woodwardin teoriasta. Luvun lopussa tarkastellaan konkreettisia esimerkkejä taloustieteen alalta. Viides luku päättää.</p> <p>Tutkielma esittää, että kausaatio tulee ymmärtää kuvauksena manipulaatioiden vaikutuksesta: kahden tapahtuman välillä on kausaalinen suhde, jos tiettyjen ehtojen vallitessa yhtä manipuloimalla voidaan vaikuttaa toiseen. Selitykset puolestaan rakentuvat näin määriteltyjen kausaalisten suhteiden varaan: ne kertovat, miten selitettävä tapahtuma riippuu johdonmukaisesti selittävästä syystä kertoen samalla, miten selitettävään tapahtumaan olisi voinut vaikuttaa. Tutkielman keskeinen havainto on, että taloustieteelliset mallit pyrkivät kuvaamaan keskeisiä selitettävään tapahtumaan vaikuttavia kausaalisuhteita, ja täten luomaan selityksiä manipulaatioteoriaa vastaavalla tavalla. Manipulaatioteorian mukaiset intuitiot vaikuttavat implisiittisesti taloustieteen metodologian taustalla, ja taloustieteellisen mallintamisen tulkitseminen tässä viitekehyksessä auttaa tarkentamaan kausaation käsitettä ja kausaaliselitysten merkitystä taloustieteessä.</p>			
Avainsanat - Nyckelord Kausaatio, tieteellinen selittäminen, taloustieteen filosofia			
Säilytyspaikka - Förvaringsställe - Where deposited Helsingin yliopiston kirjasto			
Muita tietoja - Övriga uppgifter - Additional information			

SISÄLLYS

TIIVISTELMÄ

1 Johdanto.....	1
1.1 Tieteellinen selittäminen	2
1.2 Ymmärtäminen.....	3
1.3 Kausaaliselitykset.....	5
1.4 Mallintaminen	6
1.5 Tutkielman rakenne.....	8
2 Kausaation manipulaatioteoria	9
2.1 Manipulaatiot	11
2.2 Kausaalisten suhteiden kuvaaminen	12
2.3 Manipulaatiosta kausaatioon.....	14
2.4 Interventiot	15
2.5 Syiden monimuotoisuus	18
2.6 Muuttujien valinta ja vakavasti otettavat mahdollisuudet.....	22
2.7 Manipulaatioteoria	24
3 Kausaaliselitykset.....	26
3.1 Historiallinen katsaus ja ei-kausaliset selitysteoriat.....	26
3.1.1 DN/IS-mallit.....	27
3.1.2 Puuttuva kausaatio	30
3.1.3 Unifikaatio.....	32
3.1.4 SR-malli	34
3.2 DN-mallista manipulaatioihin.....	37
3.3 Selityksellinen voima	39
3.4 Selitys ja kontrafaktuaalit.....	41
3.5 Selityksiä ilman lakeja	42

3.6 Invarianssi	44
3.7 Edistysaskeleita	46
3.7.1 Unifikaatio.....	46
3.7.2 SR-malli ja todennäköisyydet	48
3.7.3 CM-malli	50
4 Selitys taloustieteessä	52
4.1 Mallit sijaissysteeminä	54
4.1.1 Hotellingin malli	55
4.1.2 Kausaalisten tekijöiden eristäminen.....	56
4.1.3 Teoreettiset eristykset.....	58
4.1.4 Mallit ovat kokeita ja kokeet ovat malleja	60
4.1.5 Sijaissysteemien kontrafaktuaalinen uskottavuus	61
4.2 Manipulaatioteoria ja taloustieteellinen mallinnus	63
4.2.1 Mahdolliset mekanismit ja syiden monimuotoisuus	65
4.2.2 Interventiot ja eristykset.....	65
4.2.3 Invarianssi ja uskottavuus	67
4.3 Esimerkkejä taloustieteestä	68
4.3.1 Regressioyhtälöt.....	68
4.3.2 Kausaalinen päättely	71
4.3.3 Tasapainoselitykset	76
5 Pohdinta	78
Lähteet.....	82

1 Johdanto

Arkemme on täynnä selityksiä. Konkreettisimmin tämän huomaavat pienten lasten vanhemmat, jotka saavat päivittäin kymmeniä muistutuksia siitä, miten keskeinen selityksen rooli on ympäröivän maailman ymmärtämisessä ja siitä oppimisesta. Sama toistuu aikuistenkin maailmassa, vain hieman harvemmin: selitystä kaipaavat enää ainoastaan odottamattomat tapahtumat – myöhästyminen työpaikalta, outo ääni auton moottorissa, tai vanha kaveri, joka ei enää tervehtinytkään ohi kulkiessaan. Kun esimies myöhemmin tivaa selitystä myöhästymiselle, huomaamme antavamme syitä tapahtumalle: auto oli hajonnut, tai tiellä oli tapahtunut onnettomuus. Usein selityksen antaminen on syiden antamista. Tämä ajatus on meille niin luonnollinen, että se piilottaa alleen, miten hienovaraisista asioista on kyse, kun puhumme selittämisestä ja syistä – ja kausaatiosta.

Arjessa selitykset ovat joustavia, ja selityksen vaatima tarkkuus verrannollinen selitystä kaipaavaan tilanteeseen. Pienten lasten vanhemmat saavat päivittäin muistutuksia myös tästä selityksen hyväksymisen relativistisesta kriteeristä, kuten myös niiden yleisestä luonteesta: selitykset ovat vastauksia *miksi*-kysymyksiin, ja täten selitykseksi kelpaa mikä tahansa vastaus, jonka kysymyksen esittäjä hyväksyy. Siirtyessämme tieteellisen selittämisen maailmaan, intuitioidemme epätarkkuus muodostuu ongelmaksi: mikä on hyväksyttävä selitys, ja mitä ehtoja selityksemme tulee täyttää? Jos selittäminen on syiden antamista, mitä tarkoittaa, että jokin asia on toisen syy; minkä ehtojen tulee täytyä, jotta kausaalinen suhde on olemassa? Nämä kysymykset vaativat eksplisiittisiä vastauksia, sillä kuten tulemme huomaamaan, niiden vastaukset ohjaavat merkittävästi teorioiden muodostamista ja tieteellistä testaamista.

Tämän tutkielman tarkoitus on tutustua yhteen tällaiseen vastaukseen. Tarkemmin ilmaistuna tavoitteena on saavuttaa yksi mahdollinen vastaus seuraaviin kysymyksiin: 1) mitä tarkoitamme kausaatiolla, ja 2) miten tästä kausaation käsitteestä voidaan johtaa selityksiä. Tässä tutkielmassa huomio keskitetään James Woodwardin kausaaliselityksen manipulaatioteoriaan. Woodwardin mukaan selitykset ovat kuvauksia kontrafaktuaalisista suhteista tapahtumien välillä, jotka kertovat, miten selitettävä

tapahtuma seuraa johdonmukaisesti selittävistä tekijöistä. Samalla tämän kuvauksen kautta ne kertovat, miten selitettävään tapahtumaan voidaan vaikuttaa. Tämä näkemys on erittäin kytkeytynyt konkreettiseen tieteelliseen toimintaan ja testaamiseen, minkä takia tutkielman toinen tavoite koskee esitetyn teorian soveltamista yhdellä erityisalalla, taloustieteessä. Tarkoitus on luoda johdonmukainen ja kattava kuvaus siitä, kuinka Woodwardin ajatuksia voidaan hyödyntää ja tulkita taloustieteen viitekehyksessä. Yhtymäpinnan Woodwardin teorian ja taloustieteellisen metodologian välillä muodostaa Uskali Mäen eristyksiin perustuva näkemys malleista sijaisysteemeinä.

Tutkielman tarkempi rakenne esitetään alaluvussa 1.5. Ennen tätä on kuitenkin aiheellista kiinnittää huomio tiettyihin rajauksiin ja lähtökohtiin.

1.1 Tieteellinen selittäminen

Puhe tieteellisestä selittämisestä nostaa esiin kaksi jaottelua: on *tieteellistä* selittämistä, eli kaikki selitykset eivät ole tieteellisiä, ja on tieteellistä *selittämistä*, eli kaikki tieteellinen toiminta ei ole ainoastaan selittävää (Woodward, 2017). Ensimmäinen jaottelu korostaa tässäkin työssä tehtävää tärkeää rajausta juuri *tieteellisiin* selityksiin. Arjessa selitykset saavat mitä moninaisimpia muotoja: voimme selittää esimerkiksi mitä tapahtui eilen, miksi teimme jonkin tietyn ratkaisun, tai miten lempikakkumme leivotaan. Nämä eivät selvästikään ole tieteellisiä selityksiä. Tieteelliset selitykset koostuvat hyvin tietystä miksi-kysymysten osajoukosta: ne vastaavat kysymyksiin siitä, miksi jokin asia tapahtui, viittaamalla relevantteihin empiirisiin seikkoihin. Tämän kaltaisiin kysymyksiin keskittyminen rajaa käsittelyn ulkopuolelle monien arjen selitysten lisäksi myös niin kutsutut teleologiset selitykset, jotka annetaan viittaamalla selitettävän asian päämäärään tai tavoitteeseen. Tämän rajauksen kanssa on oltava tarkka, sillä useat teleologiset selitykset ovat kielellisesti identtisiä tässä työssä tavoitellun tieteellisen selittämisen kanssa, mutta käsittely on vain rajallisesti sovellettavissa näihin tapauksiin. Tässä työssä ei oteta kantaa teleologisiin selityksiin.

Tieteellisestä *selittämisestä* puhuminen valaisee puolestaan tieteellisen tiedon kaksijakoista luonnetta: meillä on tietoa *että* ja tietoa *miksi*. Ensimmäinen näistä on luonteeltaan deskriptiivistä ja toinen selittävää. (Salmon, 1989, s. 3.) Deskriptiivinen tieto

pitää sisällään laajemmin ajateltuna myös ennusteet: kun tiedämme miten asiat ovat nyt, pystymme päättelemään, miten ne tulevat kohta olemaan. Tässä kehyksessä tieteellinen selittäminen – tietoa miksi – antaa selityksen olemassa olevalle deskriptiiviselle tiedollemme. Tämän myötä tieteellinen toiminta pohjautuu kahteen erilaiseen kysymyksenasetteluun: on selitystä etsiviä kysymyksiä (”miksi X tapahtui?”) ja evidenssiä etsiviä kysymyksiä (”miksi pitäisi uskoa, että X tapahtui?”). Esimerkkejä on lukuisia. Kosmologit uskovat, että kaukaiset galaksit loittonevat meistä suurella nopeudella. Tämä uskomus perustellaan evidenssiä etsivän kysymyksen avulla, jonka vastaus pohjautuu deskriptiiviseen havaintoon galaksien lähettämän valon punasiirtymästä. Punasiirtymä ei kuitenkaan selitä, miksi galaksit loittonevat – selityksen antaa alkuräjähdyksen, jonka valtavan energian vaikutuksesta galaksit edelleen jatkavat matkaansa erilleen toisistaan. (Salmon, 1989, ss. 6–7.)

Tieteellinen selitys antaa vastauksen sellaiseen selitystä etsivään miksi-kysymykseen, joka koskee yleisesti tieteen asiasisältöä. Tämä on karkeasti myös se rajaus, jota tässä työssä käytetään. Useat arkielämän selitykset ovat samankaltaisia tieteessä käytettyjen selitysten kanssa, joten samoja käsitteitä voidaan soveltaa myös monien arkielämän esimerkkien kohdalla. Näin tullaan myös tässä työssä tekemään sen ollessa aiheellista. Rajauksen tarkoitus on kuitenkin eksplisiittisesti sulkea käsittelyn ulkopuolelle sellaiset miksi-kysymykset, joihin teoriat eivät pysty, tai edes yritä, antaa vastausta. Käsitellessämme tieteellistä selittämistä on hyväksyttävää, että käsittely soveltuu vain rajattuun osaan moninaisista arkemme selityksistä. Jatkossa kun puhumme selittämisestä, tarkoitamme näin rajattua tieteellistä selittämistä.

1.2 Ymmärtäminen

1900-luvun alkupuolella voimissaan oli näkemys, jonka mukaan tiede on rajoittunut ainoastaan havaittaviin ilmiöihin. Tällöin ajateltiin, että empiirinen tiede ei pysty vastaamaan selitystä etsiviin kysymyksiin, sillä tieteellinen tieto on luonteeltaan ainoastaan deskriptiivistä. Jos haluamme tietää, miksi jokin asia tapahtui, meidän on siirryttävä empiirisen tieteen ulkopuolelle ja vedottava esimerkiksi metafysiikkaan. (Salmon, 1998, ss. 79–80.) Empiirinen tiede ei pysty yksin selittämään mitään, koska

yhteyden luominen kahden tapahtuman välille puhtaasti havainnoimalla ei ole mahdollista – voimme ainoastaan havaita tapahtumia ja niiden seurauksia, mutta se, miksi juuri näillä tapahtumilla on juuri nämä seuraukset, on tieteen tavoittamattomissa. Tämä näkemys on nykyajan valossa vähintäänkin liian ankara, sillä kuten tulemme huomaamaan, kausaatio voidaan määritellä sopimaan erinomaisesti yhteen empiirisen käytännön kanssa.

Viimeisen vuosisadan aikana ymmärryksemme maailman toiminnasta on lisääntynyt merkittävästi. Tällä on suora yhteys tieteen kehitykseen, joka on mahdollistanut selitysten antamisen yhä kasvavalle joukolle tapahtumia – selityksillä pyritään ymmärrykseen, ja ymmärrys muodostuu hyväksyttävästä ja sisäistetystä selityksestä. Tämä päämäärä on tärkeä pitää mielessä, sillä se tarjoaa yhden olennaisen ulottuvuuden kilpailevien selitysten arvioinnille; ymmärryksellä on merkittävää käytännöllistä ja älyllistä arvoa, jota pelkän selityksen olemassaolo ei pysty tavoittamaan. Täten selityksiä voidaan arvioida sen suhteen, miten ne onnistuvat ymmärryksen tuottamisessa. Tätä kuvataan usein puhumalla selityksen *voimasta* tai *syvyydestä*.¹ Yksi tämän työn tavoitteista on osoittaa, että Woodwardin selitysteoria tarjoaa paremman pohjan ymmärryksen muodostumiselle kuin kilpailevat historialliset teorit – sen tarjoamat selitykset ovat parempia ja voimakkaampia.

Meidän tulee pyrkiä kohti selityksiä, koska ne lisäävät ymmärrystämme. Koska sana *ymmärrys* tulee välttämättä nousemaan esiin myös jatkossa, on sen kohdalla hyvä tehdä samanlainen erottelu kuin selityksen kohdalla aiemmin. Arjen ymmärrys koostuu monista erilaisista asioista, joihin kuuluvat muun muassa merkitykset, toisten ihmisten tunteet ja ajatukset, aikomukset ja tarkoitukset. Nämä toimivat selityksinä monille arkielämän tapahtumille, ja siten ne myös luovat ymmärrystä meitä ympäröivistä asioista. Kuten selitysten kohdalla, rajaamme tässä työssä myös ymmärryksen käsitteen koskemaan tieteellisen selityksen luomaa ymmärrystä tieteen asiasisältöön kuuluvista asioista.

¹ Ylikoski ja Kuorikoski (2010) hahmottavat viisi selityksellisen voiman ulottuvuutta, joihin palataan tarkemmin alaluvussa 3.3.

1.3 Kausaaliselitykset

Suuri osa tästä työstä koskee Woodwardin teoriaa kausaaliselityksistä, joten on aiheellista esittää muutama kommentti tästä filosofisesti mielenkiintoisesta rajauksesta *kausaalisiin* selityksiin. Tässä vaiheessa tyydytään intuitiiviseen käsitykseen kausaalisesta selittämisestä: tapahtuman selitys on kausaalinen, jos se annetaan viittaamalla kyseisen tapahtuman syihin. Rajaus sisältää oletuksen, että kausaaliset selitykset muodostavat vain osan kaikista mahdollisista selityksistä, olkoonkin tämä osa kuinka iso tahansa. Tämä on kiistelty aihe: joidenkin mielestä kaikki selitykset ovat kausaalisia. Helppoin tapa argumentoida tätä näkemystä vastaan on esittää hyväksyttävä selitys, joka ei ole kausaalinen. Kandidaatteja on lukuisia, kuten esimerkiksi:

- 1) Selitykset, jotka ovat totta jonkin tapahtuman syvemmän faktan nojalla (*in-virtue-of explanations*). Kuvitellaan jonkun kysyvän, miksi pisteiden A ja B välinen etäisyys on viisi metriä. Yksi mahdollinen vastaus (ja näin ollen selitys) on, että lyhin reitti pisteiden A ja B välillä on viisi metriä pitkä. (Skow, 2014, ss. 446–447.)
- 2) Selitykset, jotka viittaavat luonnonlakeihin. Esimerkiksi: ”Tuo meteori liikkuu hitaammin kuin valo, koska mikään aineellinen olio ei voi liikkua nopeammin (tai yhtä nopeasti kuin) valo (Skow, 2014, s. 455)”.
- 3) Mikrotason selitykset. Russellin (1913) kuuluisan vasta-argumentin mukaan fundamentaalifysiikassa tapahtumat selitetään funktionaalisten suhteiden kuvauksen eikä syy-seuraussuhteiden kautta. Täten mikrotasolla ei ole olemassa kausaatiota siinä mielessä kuin sen makromaailmassa ymmärrämme, ja sen selitykset eivät ole kausaalisia.
- 4) Konstitutiiviset selitykset. Kysymykset ”Miten lasista tuli särkyvä?” ja ”Miksi lasi on särkyvä?” vaativat erilaisen vastauksen, vaikka ne voidaan käytännössä nähdä identtisinä. Ensimmäiseen kysymykseen voidaan vastata kuvaamalla särkymisen mahdollisuuden johtanut kausaalinen historia: mitkä asiat lasin syntyprosessissa aiheuttavat sen särkyvyyden. Toinen kysymys puolestaan vaatii kuvauksen niistä lasin ominaisuuksista, jotka tekevät särkymisen mahdolliseksi – tämä selitys *ei* ole

kausaalinen vaan konstitutiivinen, sillä se koskee lasin rakennetta eikä sitä, miten se on nämä ominaisuudet saanut. (Ylikoski, 2013, ss. 279–280.)

- 5) Tasapainoselitykset. Tasapainoselitysten tarkoitus ei ole kuvata kausaalisia tekijöitä, vaan osoittaa, että tasapainoon olisi päädytty riippumatta systeemin alkutilasta. Tasapainoon johtavat mikrotason kausaaliset tekijät ovat selityksen kannalta irrelevantteja. (Sober, 1983, s. 207.)

Jotta kaikki selitys voitaisiin nähdä kausaalisenä, olisi pystyttävä osoittamaan, että vastaesimerkit eivät joko ole selityksiä lainkaan, tai että ne ovat tosiasianssa kausaalisia. Woodward ei ryhdy tähän projektiin ja tyytyy toteamaan, että hänen käsittelynsä on sovellettavissa ainoastaan kausaalisiin selityksiin: mahdollisuus ei-kausaalisten selitysten olemassaololle jätetään auki. Lopullinen rajanveto riippuu siitä, miten määrittelemme kausaation ja selityksen, sekä niiden välisen suhteen. Tämäkään ei ole yksiselitteistä: Skow (2014, ss. 456–457) näkee lakeihin viittaavat selitykset kausaalisenä, ja hänen mukaansa ne voidaan nähdä kausaalisenä jopa Woodwardin mallin sisällä. Woodward (2003, ss. 208–209) on tästä eri mieltä. Tasapainoselitykset Woodward (2003, s. 6) puolestaan näkee kausaalisenä selityksinä – tähän ajatukseen palaamme tarkemmin taloustiedettä koskevassa luvussa.

Mahdollisten vastaesimerkkien on tarkoitus tässä vaiheessa osoittaa, että kausaaliselityksen teorioiden sovellusalaan on aiheellista kiinnittää huomiota, ja että seuraavien lukujen käsittely ei välttämättä ole sovellettavissa kaikkeen tieteelliseen selittämiseen. Woodwardin manipulaatioteoria tarjoaa kuitenkin hyvin kattavan esityksen kausaaliselityksistä, jotka epäilemättä muodostavat suuren osan tieteellisistä selityksistä. Manipulaatioteorian avulla voidaan näin ollen käsitellä laajaa tieteellisen toiminnan skaalaa niin luonnon- kuin sosiaalitieteiden saralla.

1.4 Mallintaminen

Tieteenfilosofisesti yksi suurimmista murroksista sosiaalitieteissä (ja näin ollen myös taloustieteessä) on koskenut mallinnuksen kasvanutta merkitystä selitysten luomisessa. Historiallisesti tieteellinen toiminta on keskittynyt teorioihin, joiden sisältämien väitteiden nähtiin olevan suoraan yhdistettävissä empiirisiin havaintoihin. Tässä

näkemyksessä selitykset muodostuvat tietyn tapahtuman johtamisesta yleisemmästä (todellisuutta kuvaavasta) teoriasta. Teoreettiset väitteet ovat kuitenkin harvoin näin helposti yhdistettävissä havaintoihin. Tästä syystä sekä tieteentekijöiden että filosofien huomio alkoi 1900-luvun loppupuolella yhä enemmän kohdistua malleihin. Ajatus on, että suoraan todellisuutta kuvaavan teorian sijaan voimme luoda todellisuudesta erillisen mallin, joka toimii teorian käytännöllisenä kuvauksena. Mallia hyödyntämällä ja tarkastelemalla voimme kuitenkin tuottaa tietoa ja selityksiä sen kuvaamasta aidosta tapahtumasta. (Kuorikoski, 2010b, ss. 12–14.)

Taloustieteessä mallit ovat läsnä kaikessa toiminnassa, ja niihin suhtaudutaan – tai ainakin tulisi suhtautua – lähes täysin pragmaattisesti. Mallien tavoite on yksinkertaisesti tavoittaa todellisuudesta kaikkein relevantteimmat osa-alueet tarkasteltavan tapahtuman kannalta. Tämän saavuttaminen vaatii malleilta kykyä mukautua: yhteen tapahtumaan soveltuva malli ei välttämättä sosiaalisen toiminnan monimuotoisuuden takia ole lainkaan käyttökelpoinen toisessa tilanteessa. Mielipide-erot relevanttien seikkojen valinnasta johtavat usein myös tilanteisiin, joissa useampi malli antaa täysin erilaisen selityksen samalle tapahtumalle, ilman että mikään selityksistä olisi yksiselitteisesti oikea tai väärä. Yksi mallien tärkeimmistä tehtävistä on kuvata sitä, mitkä mekanismit tapahtumaan vaikuttavat (tai voisivat vaikuttaa)², ja nämä voivat vaihdella merkittävästi pieniltä vaikuttavien asioiden ja oletusten myötä. Mallintamisen hyödyllisyys ei varsinaisesti riipu siitä, kertovatko ne tarkasti millaisiin lopputuloksiin tietystä alkutilasta päädytään, vaan enemmän siitä, että ne paljastavat, mistä tekijöistä todennäköinen lopputulos riippuu.

Monimuotoisuuden lisäksi keskeinen mallinnuksen ominaisuus on pyrkimys yksinkertaisuuteen. Mallien selityksellinen voima kumpuaa niiden kyvystä pelkistää tapahtuman kannalta relevantit asiat johdonmukaiseen ja helposti ymmärrettävään muotoon. Dani Rodrik (2015, s. 43) kuvaa yksinkertaisuuden tavoitteen hyvin osuvasti Borgesin novelliin perustuvalla esimerkillä, jossa muinaisen valtakunnan kartantekijät halusivat kartoistaan yhä tarkempia, ja päätyivät lopulta sisällyttämään karttoihinsa koko maanpinnan 1:1 mittakaavassa siten, että kartta peitti koko maaston. Tällaisesta kartasta

² Mallien hierarkiassa tämä tehtävä yhdistyy teoreettisiin malleihin (ks. Suppes, 1962).

ei ole enää mitään hyötyä navigoinnin kannalta – liika tarkkuus ja monimutkaisuus kadottaa pelkistetyn mallin selityksellisen arvon. Yksinkertaisuudesta kuitenkin seuraa välttämättä se, että mallit eivät koskaan ole *tosia* sanan tiukassa merkityksessä. Mikäli haluamme johtaa malleista todellisuutta koskevia päätelmiä, meidän on luotava jonkinlainen yhteys näiden välille – miten voimme ”epätarkoista” ja yksinkertaisista malleista johtaa relevanttia informaatiota monimutkaisista aidoista tapahtumista?

Yksi vastaus tähän selitysparadoksinakin tunnettuun ongelmaan (ks. Reiss, 2012) pohjautuu mallien ja empiiristen kokeiden yhteyksiä korostavaan näkemykseen mallinnuksesta, jota tässä työssä edustaa luvussa 4 esiteltävä Uskali Mäen teoria. Yksi tämän työn keskeisistä havainnoista on Woodwardin teorian yhteensopivuus empiirisen käytännön kanssa, ja vastaavasti se tarjoaa erinomaisen pohjan taloustieteellisten mallien kausaalille tulkinnalle (ja täten myös selittävyydelle), siinä missä nämä mallit mukailevat samoja käytäntöjä.

1.5 Tutkielman rakenne

Tämän tutkielman tavoite on antaa yksi mahdollinen tulkinta kausaatiosta, kausaaliselityksistä, ja näiden yhteydestä taloustieteen mallintamisen metodologiaan. Tutkielma etenee seuraavasti.

Toisessa luvussa esittelen Woodwardin näkemyksen kausaatiosta, joka luo käsitteellisen pohjan selitysten ymmärtämiselle tässä viitekehyksessä. Woodwardin keskeinen havainto on, että pelkkä syistä ja seurauksista puhuminen on harhaanjohtavaa, ja että käsitteistöä on laajennettava, jotta kausaaliset väitteemme saavuttavat tarkemmin kuvattavan tapahtuman todellisen kausaalisen rakenteen. Tapahtumien pohjalla on aina olemassa jokin kausaalinen rakenne, jonka voimme selvittää tarkastelemalla tapahtuman kannalta relevantteja manipulaatioita ja niiden seurauksia – kahden tapahtuman välillä on kausaalinen suhde, jos yhtä manipuloimalla voidaan vaikuttaa toiseen. Toisen luvun on tarkoitus antaa määritelmä siitä, mitä Woodwardin teoriassa kausaaliset väitteet todellisuudessa merkitsevät, ja minkä ehtojen vallitessa voimme sanoa, että kahden asian välillä on kausaalinen suhde.

Kolmannessa luvussa huomio kiinnitetään selityksiin. Luku alkaa historiallisella katsauksella, jossa esitetään Woodwardin teorian kannalta kolme historiallisesti tärkeintä selitysteoriaa: DN/IS-malli, unifikaatioteoria ja SR-malli. Näitä hyödyntäen pyritään perustelemaan, että kausaation käsite on olennainen selitysten kannalta. Näitä ajatuksia jatketaan esittämällä Woodwardin kausaatioon perustuva selitysteoria kontrastina historiallisille teorioille. Tavoitteena on yleisemmin osoittaa, että Woodwardin teoria tuottaa parempia selityksiä, mutta myös korostaa niitä tekijöitä, jotka tekevät Woodwardin mallista erittäin käyttökelpoisen sosiaalitieteiden parissa. Kuten ensimmäisen luvun kohdalla, toisen luvun on tarkoitus määrittellä, mitä selittäminen Woodwardin teoriassa tarkoittaa, ja minkälaisen seikkojen kuvauksesta selitys muodostuu. Luvun lopussa pyritään lyhyesti kuvaamaan, kuinka Woodwardin malli onnistuu sisäistämään useita vahvuuksia edeltäneistä teorioista, joista on hyötyä myös tulevan käsittelyn kannalta.

Neljäs luku keskittyy taloustieteeseen, ja taloustieteellisen mallintamisen ja kausaalisen selittämisen yhteyksiin. Esittelen aluksi Uskali Mäen teorian, jonka mukaan mallit ovat sijaissysteemejä, joita tarkastelemalla voimme saada tietoa niiden kuvaamista aidoista tapahtumista. Tämän jälkeen argumentoidaan, että useat tämän teorian keskeiset käsitteet ovat taloustieteeseen sopivia sovelluksia Woodwardin teorian ajatuksista. Tavoitteenani on esittää Woodwardin teoriaan pohjautuva kausaalinen luenta taloustieteellisestä mallintamisesta, niin kuin se, ja sen metodit ja päämäärä, Mäen teoriaa mukaillen ymmärretään. Luvun lopussa esitetään konkreettisia esimerkkejä tällaisen käsittelyn soveltamisesta taloustieteen malleihin ja metodeihin. Viides luku päättää.

2 Kausaation manipulaatioteoria

Kausaatio on tärkeää – tämä ajatus sisältyy jo tilastollisen työskentelyn kenties toistetuimpaan mantraan: korrelaatio ei ole kausaatiota. Tilastollisten keinojen avulla voidaan helposti havaita riippuvuussuhteita muuttujien välillä, mutta kausaaliin väitteisiin suhtaudutaan aina varovasti. Vaikeudestaan huolimatta ajatus kausaatiosta pohjustaa kuitenkin lähes kaikkea sosiaalitieteiden tutkimusta. Korrelaatiot ovat mielenkiintoisia, mutta ilman kuvausta muuttujien välisistä kausaalisuhteista,

teorioitamme puuttuu selitysvoima. Jos olemme huolissamme kotitalouksien velkaantumisesta, ja haluamme tehdä asialle jotain, se mitä tapahtuu yhtäaikaisesti velkaantumisen kanssa, on kysymyksen kannalta irrelevanttia – mitä me todellisuudessa haluamme tietää ovat ne syyt, joista velkaantuminen johtuu.

James Woodward esittää teoksessaan *Making Things Happen* (2003) manipulaatioihin perustuvan näkemyksensä kausaatiosta. Woodwardin tavoitteena on luoda laaja-alainen kuvaus kausaaliselityksistä, niin kuin ne ilmenevät monissa sekä tieteen että arkielämän konteksteissa. Woodward näkee kausaaliselitysten olevan sellaisten selitysten luokka, joka kertoo riippuvuussuhteista lopputulosten ja niiden aiheuttajien välillä. Tällaisille selityksille on Woodwardin kannalta lisäksi ominaista se, että ne ovat relevantteja tapahtumien kontrolloinnin ja manipulaation kannalta: kausaaliselitykset kertovat meille, miten tutkittava tapahtuma reagoi mahdollisiin muutoksiin siihen vaikuttavissa tekijöissä. (Woodward, 2003, s. 6.) Kausaalinen suhde tapahtumien välillä tarkoittaa sitä, että yhtä manipuloimalla voidaan vaikuttaa toiseen.

Kausaatiolla on Woodwardin mallissa suuri käytännöllinen merkitys. Vain kausaalisen tiedon avulla voimme tehdä hyviä päätöksiä pyrkiessämme tiettyihin lopputuloksiin. Henkivakuutus on tästä hyvä esimerkki: on fakta, että henkivakuutusten ottajat elävät kauemmin, mutta tästä huolimatta henkivakuutuksen ostaminen ei ole hyvä strategia oman eliniän kasvattamiseksi. Tämä kuvaa tyypillistä käytännön tilannetta, jossa voimme havaita kahden asian olevan korreloituneita, mutta jossa niiden välillä ei ole kausaatiota. Voimme havainnoimalla tulla päätelmään eliniän ja henkivakuutuksen yhteydestä, mutta kausaation kannalta keskeinen kysymys on se, voiko yksittäinen ihminen vaikuttaa omaan elinikäänsä ostamalla vakuutuksen. Vastaus tässä tapauksessa on ilmeinen, mutta sen varmistaminen vaatisi tosiasiaissa satunnaiskokeen toteuttamista. Odotusten mukaisesti tällaisessa kokeessa korrelaatio katoaisi ja voimme huomata, että henkivakuutuksen ja eliniän välillä ei ole kausaalista suhdetta: toista manipuloimalla ei voida vaikuttaa toiseen. (Woodward, 2003, ss. 28–32.)

Yksi syy Woodwardin teorian menestykseen on epäilemättä sen käytännön tieteelliseen toimintaan soveltuva näkemys kausaatiosta. Useiden ekonometrian oppikirjojen

määritelmät antavat tästä todisteita: Stock ja Watson (2020, s. 48) mieltävät kausaalisen vaikutuksen olevan ”tietyn tapahtuman vaikutus lopputulokseen, kuten se on mitattu ideaalissa kontrolloidussa satunnaiskokeessa”. Samankaltaisen tulkinnan esittää myös Verbeek (2017, s. 63) puhuessaan regressioyhtälöiden kertoimista (jotka kertovat miten suuri vaikutus tietyllä muuttujalla on odotettuun lopputulokseen): ”jos haluamme tulkita kertoimet kausaalisesti, *ceteris paribus*-ehdon tulee sisältää myös kaikki muut (havaitut ja ei-havaitut) muuttujat”. *Ceteris paribus* voidaan suomentaa ”muiden asioiden pysyessä yhtäläisenä”, ja ehto takaa, että ideaalin kontrolloidun satunnaiskokeen ehdot toteutuvat: kaikki muut paitsi tarkasteltava muuttuja kontrolloidaan muuttumattomiksi, ja jäljelle jäävä (vain tarkasteltavan muuttujan) vaikutus lopputulokseen voidaan mieltää kausaaliseksi vaikutukseksi.

Ekonomistit harvoin perustelevat valitsemaansa kantaa filosofien vaatimalla perusteellisuudella, mikä voi helposti johtaa käytäntöjen hylkäämiseen pinnallisina. Tämä on kuitenkin Woodwardin (2003, ss. 26–27) mukaan virhe, sillä niiden menestys paljastaa, että ne ovat vähintään tutkimisen arvoisia. Woodward pyrkii täten pelastamaan tieteellistä toimintaa lähellä olevat näkemykset ja tarjoamaan niille vakavasti otettavan filosofisen pohjan. Siirrytään seuraavaksi tutkimaan mallia yksityiskohtaisemmin.

2.1 Manipulaatiot

Manipulaatioteoriassa kausaalisuhteet nähdään suhteina, joissa kaksi tai useampi muuttuja liitetään toisiinsa siten, että kun yhden arvo muuttuu, se muuttaa myös muiden muuttujien arvoja.³ Muuttujat voivat olla ominaisuuksia tai määriä, ja ne voivat saada binäärisiä tai jatkuvia arvoja (tai mitä tahansa tältä väliltä) (Woodward, 2003, s. 40.) Kun sanotaan, että X on Y:n syy, se tarkoittaa manipulaatioteorian mukaan vain, että on olemassa *jokin* X:n manipulaatio, jolla voidaan saada aikaan *jokin* muutos Y:ssä tai sen todennäköisyysjakaumassa. Kontrafaktuaalisessa mielessä kausaalisen suhteen voidaan

³ Huomioitavaa on, että muuttaminen on myös kausaalinen käsite, joka olettaa kausaalisen suhteen muutoksen aiheuttajan ja muuttuvan tekijän välillä. Tästä seuraavaa mahdollista kehäisyyttä käsitellään tarkemmin interventioita koskevassa luvussa 2.4.

ymmärtää kertovan meille, miten Y reagoisi, *jos* tietty muutos X:ssä tapahtuisi.⁴ Tällöin voimme teorian avulla ymmärtää kausaalisesti myös väitteitä, kuten ”tupakointi aiheuttaa keuhkosityöpää”. Väite sellaisenaan ei kerro mitään yhdenkään yksittäisen henkilön keuhkosityövän ilmaantumisen tai parantumisen, eikä se edes vaadi, että kukaan maailmassa oikeasti tupakoisi. Olennaista on vain, että mikäli jonkin agentin kohdalla tupakointia kuvaava binäärinen muuttuja saisi arvon {tupakoi}, se vaikuttaisi keuhkosityöpää kuvaavaan muuttujaan siten, että tämän muuttujan todennäköisyys saada arvo {sairastuu keuhkosityöpään} kasvaisi. (Woodward, 2003, s. 40.)

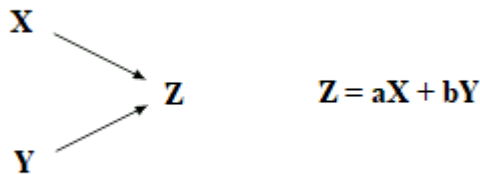
Woodward käyttää väitteitä ”X on Y:n syy” ja ”X on kausaalisesti relevantti Y:n kannalta” synonyymisinä, mikä eroaa hieman siitä, miten arkielämässä ymmärrämme tällaiset väitteet. Arjessa väite ”X aiheuttaa Y:n” ymmärretään usein niin, että se yhdistää muuttujan arvoja itse muuttujien sijaan. Toisin sanoen, esimerkiksi tupakoinnin ja keuhkosityövän tapauksessa väite yhdistää tupakoinnin (X:n arvon {tupakoi}) ja keuhkosityövän (Y:n arvon {sairastuu keuhkosityöpään}) sen sijaan, että se yhdistäisi muuttujat toisiinsa; arjessa tällaiset väitteet ymmärretään usein ainoastaan positiivisen vaikutuksen kautta. Väite ”X on kausaalisesti relevantti Y:n kannalta” puolestaan kattaa laajemman skaalan tapahtumia, joissa X on joko positiivisesti tai negatiivisesti relevantti Y:n tapahtumisen kannalta. Jatkossa kun puhutaan syistä, niillä tarkoitetaan tätä laajempaa tulkintaa, joka kattaa myös negatiiviset relevanssit: tupakointi on relevanttia myös keuhkosityövän puuttumisen kannalta, vaikka tupakointi itsessään ei tätä aiheuta. (Woodward, 2003, s. 41.) Yhteys on tärkeää luoda muuttujien eikä niiden saamien arvojen välille: olennaista on se, että X:n arvon manipulointi vaikuttaa Y:n arvoon tai sen todennäköisyysjakaumaan.

2.2 Kausaalisten suhteiden kuvaaminen

Kuvatessaan kausaalisia suhteita ja niiden rakenteita Woodward käyttää kahta työkalua: graafeja ja funktioita. Graafi on järjestetty pari, jossa kausaalisen rakenteen kaksi osaa

⁴ Muuttujia yhdistävät kausaaliset väitteet tunnetaan kirjallisuudessa tyyppikausalisina (*type-causal*) väitteinä, jotka eroavat tietyiltä osin aitoja yksittäisiä tapahtumia yhdistävistä kausaalisista (*token-causal* tai *actual cause*) väitteistä. Tässä työssä keskitytään tyyppitasoon kausaatioon, mutta tämä käsittely on yleistettävissä myös yksittäisiin tapahtumiin (ks. Woodward, 2003, pp. 74–86).

yhdistetään toisiinsa nuolella (Woodward, 2003, s. 42). Nuoli muuttujasta X muuttujaan Y kertoo, että X on Y:n suora syy. Tämä käsite määritellään tarkemmin myöhemmin, mutta yleistäen se tarkoittaa, että X:n manipulointi vaikuttaa Y:hyn suoraan ilman välikäsiä, kun kaikki muut muuttujat on kontrolloitu muuttumattomiksi. Tämä tarkoittaa myös, että mikä tahansa vaikutus Y voidaan määritellä kaikkien sen suorien syiden funktiona.



Vasemmalla graafinen kuvaus kausaalisesta suhteesta: X ja Y ovat Z:n suoria syitä. Oikealla (mahdollinen) funktionaalinen kuvaus samasta suhteesta. (Woodward, 2003, ss. 44–45.)

Funktiot luovat deterministisen kuvauksen siitä, miten tarkasteltava muuttuja reagoi sen suorissa syissä tapahtuviin muutoksiin. Ne voidaan lisäksi nähdä kontrafaktuaalisina kuvauksina siitä, mitä tapahtuisi suorien syiden saadessa tiettyjä arvoja: asettaessamme yllä olevan kuvauksen kertoimille a ja b, ja muuttujille X ja Y tietyt arvot, Z:n saama arvo voidaan laskea suoraan funktiosta. Mikäli kaikki suorat syyt eivät ole tiedossa, funktio voidaan rakentaa perustuen kaikkiin tiedossa oleviin suoriin syihin ja virhetermiin, jonka on tarkoitus kattaa kaikkien tuntemattomien muuttujien suora vaikutus tarkasteltavaan muuttujaan. (Woodward, 2003, s. 43.)

Graafisen ja funktionaalisen esityksen suhteesta on tehtävä vielä yksi huomio. Graafinen esitys, jossa kulkee nuolet joukosta $X_1 \dots X_m$ johonkin Y:hyn, kertoo vain, että Y:n arvo voidaan esittää jonain funktiona joukosta $X_1 \dots X_m$. Tarkasti ottaen se tarkoittaa, että jokaiselle X_i on olemassa jokin yhdistelmä joukon muiden muuttujien arvoja siten, että näiden arvojen vallitessa muutos X_i :ssä aiheuttaa muutoksen Y:ssä. Graafinen esitys ei kuitenkaan täsmennä tätä enempää. Funktionaalinen kuvaus sisältää tätä enemmän informaatiota, sillä se kertoo lisäksi, *miten* Y muuttuu X_i :n muuttuessa. (Woodward, 2003, ss. 43–44.) Tästä syystä graafiset kuvaukset eivät pysty erottamaan tilanteita, joissa jokin

suora syy vaikuttaa lopputulokseen vain tiettyjen muiden suorien syiden arvojen ollessa voimassa. Woodward (2003, ss. 44–45) kuvaa tätä esimerkillä, jossa on kolme binääristä muuttujaa, jotka saavat joko arvon 0 tai 1: H kuvaa hapen saatavuutta, O kuvaa oikosulkua ja T kuvaa tulipaloa. Näiden suhde voidaan esittää funktionaalisesti: $T = H * O$. Mikäli sekä H että O saavat arvon 1, tulipalo syttyy (T saa arvon 1), mutta jos edes toinen näistä saa arvon 0, T saa myös arvon 0. Koska on olemassa *jokin* manipulaatio, jolla sekä H että O vaikuttavat T:n arvoon (silloin kun toinen saa arvon 1), molemmat ovat T:n suoria syitä, ja graafinen kuvaus on täysin samanlainen kuin edellä: kaksi nuolta T:hen, toinen H:sta ja toinen O:sta. Graafinen kuvaus ei kuitenkaan riitä tunnistamaan tällaisia tilanteita, jossa muuttujien välillä on vuorovaikutusta.

Jatkossa kuvaajat ja funktiot ymmärretään tässä esitetyn kaltaisina kuvauksina kausaalisista suhteista muuttujien välillä.

2.3 Manipulaatiosta kausaatioon

Kausaation sitominen manipulaatioon vaatii määritelmän siitä, miten manipulaatioista ja niiden seurauksista voidaan tehdä johtopäätöksiä kausaalisista suhteista. Alustavaa tarkastelua varten Woodward esittää kausaatiolle riittävän ehdon (**RE**) seuraavasti:

”(**RE**) Jos (i) on olemassa mahdollinen interventio, joka muuttaa X:n arvoa siten, että (ii) tämän (ja vain tämän) intervention toteuttaminen muuttaa Y:n arvoa, tai Y:n todennäköisyysjakaumaa, niin X on Y:n syy (Woodward, 2003, s. 45).”

Vastaavasti **RE**:n pohjalta määritellään välttämätön ehto (**VE**):

”(**VE**) Jos X on Y:n syy, niin (i) on olemassa mahdollinen interventio, joka muuttaa X:n arvoa siten, että (ii) jos tämä (ja vain tämä) interventio toteutettaisiin, Y:n arvo, tai Y:n todennäköisyysjakauma, muuttuisi (Woodward, 2003, s. 45).”

Määritelmiin liittyen on tehtävä ainakin kolme huomiota: ensinnäkin, niissä puhutaan interventioista. Interventio on tietty, hyväksyttävä tapa suorittaa manipulaatio – kaikki manipulaatiot eivät ole interventioita. Toiseksi, ehdot rajoittavat tehtävät interventiot vain yhteen kerrallaan. Tämän on tarkoitus sulkea pois mahdolliset tilanteet, joissa X:ään

kohdistuva interventio saa aikaan lisäksi erillisen Y:tä koskevan intervention: tällöin muutos Y:ssä on tämän toisen intervention eikä X:n muutoksen kausaalisen vaikutuksen aiheuttamaa. Kolmanneksi, ”mahdollinen interventio” on ymmärrettävä kontrafaktuaalisesti: selvää on, että emme halua sulkea pois kaikkia kausaalisia suhteita, joihin ihminen ei nykyisellään pysty vaikuttamaan. ”Mahdollinen interventio” sisältää siis laajemmin kaikki tilanteet, joissa kausaatio pohjautuu sellaiseenkin suhteeseen, jota emme nykyisellään pysty konkreettisesti manipuloimaan. (Woodward, 2003, s. 46.)

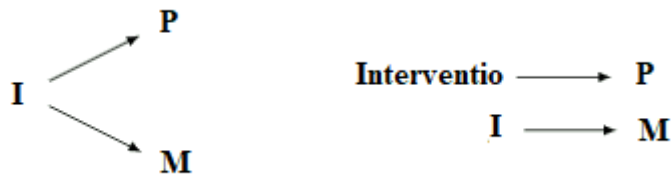
Minkälaiset X:ään kohdistuvat manipulaatiot voidaan laskea interventioiksi? Oetaan esimerkiksi kausaalinen rakenne, jossa ilmanpaine (I) aiheuttaa sekä ilmapuntarin (P) lukeman että myrskyn (M) ilmaantumisen. P:n ja M:n välillä ei tosiasiasa vallitse kausaalista suhdetta, mutta on selvää, että on olemassa tapoja muuttaa ilmapuntarin lukemaa siten, että nämä muutokset vastaavat muutoksia myrskyn ilmaantumisessa. Näin tapahtuu esimerkiksi, mikäli puntarin lukemaa muutetaan vaikuttamalla ilmanpaineeseen, tai mikäli puntariin tehtävät muutokset ovat täydellisesti korreloituneita ilmanpaineen muutoksen kanssa. (Woodward, 2003, s. 46.) Koska manipulaatiomallilla on selvä yhteys käytännön tieteelliseen testaamiseen, tällaiset tilanteet on helppo ymmärtää huonosti rakennettuina kokeina: ne pyrkivät testaamaan puntarin lukeman vaikutusta myrskyn ilmaantumiseen, mutta epäonnistuvat siinä huonon suunnittelun takia. Aiemmin ilmaistu *ceteris paribus*-ehto ei toteudu: testattava muuttuja on täydellisesti korreloitunut mallista puuttuvan muuttujan kanssa, joten päädyimme vakavasti yliarvioimaan ilmapuntarin kausaalisen vaikutuksen myrskyn kannalta.

2.4 Interventiot

Hyväksyttävät interventiot voidaan nähdä hyvin rakennettuina kokeina. Strevens (2007, s. 234) kuvaa interventioita tilanteella, jossa tutkitaan vaikuttaako kahvin juonti syövän mahdollisuuteen. On mahdollista, että kahvi itsessään kasvattaa syövän mahdollisuutta, mutta on myös mahdollista, että esimerkiksi stressi lisää yhtäaikaaisesti kahvin juontia ja syövän todennäköisyyttä. Tässä tapauksessa interventioiksi lasketaan sellaiset manipulaatiot, jotka vaikuttavat vain ja ainoastaan kahvin juomiseen, ja tämän kautta

(mahdollisesti) syövän todennäköisyyteen. Interventio on ikään kuin ”jumalan käsi”, joka täyttää tutkittavan henkilön kahvilla vaikuttamatta mihinkään muuhun tekijään.

Tarkemmin muotoiltuna intervention on täytettävä tiettyjä ehtoja. Näitä on helppo tutkia mieltämällä interventio muuttujaksi, joka voidaan nähdä osana tapahtuman kausaalista rakennetta. Mikäli haluamme tutkia, onko X Y:n suora syy, interventio määritellään suhteessa X:ään siten, että I on X:n interventiomuuttuja, jos 1) I aiheuttaa X:n, 2) I sulkee pois kaikki muut X:n suorat syyt, 3) kaikki suorat reitit I:sta Y:hyn kulkevat X:n kautta, ja 4) I on riippumaton muista muuttujista, jotka aiheuttavat Y:n. Intervention on siis oltava X:n saaman arvon aito syy eikä se saa vaikuttaa Y:hyn mitään muuta reittiä. (Woodward, 2003, s. 98.) Kahvin ja syövän tapauksessa jumalan käsi täyttää kaikki nämä ehdot, kuten tekee myös sellainen kahvin määrän kasvattaminen, joka johtuu esimerkiksi ainoastaan tutkijan vaatimuksesta – tietenkin sillä oletuksella, että lisääntynyt kahvin juominen ei johda muihin oheisvaikutuksiin, jotka vaikuttavat syövän todennäköisyyteen jotain muuta reittiä. Voimme lisäksi huomata, että aiemmat ilmapuntariesimerkkiä koskevat huonosti rakennetut kokeet eivät täytä intervention vaatimuksia: jos tehdyt muutokset ovat täydellisesti korreloituneet ilmanpaineen muutoksen kanssa, rikomme ehtoa (4), ja jos muutamme puntarin lukemaa manipuloidulla ilmanpainella, rikomme ehtoa (3). Hyvin rakennettu koe antaa puntarille satunnaisia arvoja siten, että se ei voi olla korreloitunut ilmanpaineen (tai muiden myrskyn suorien syiden) kanssa, tai pysty vaikuttamaan myrskyn ilmaantumiseen mitään muuta reittiä. Tällaisessa testissä aiemmin havaittu yhteys myrskyn ja puntarin välillä katoaa.



Vasemmalla puolella olemassa oleva kausaalinen rakenne ilmanpaineen (I), puntarin (P) ja myrskyn (M) välillä. Oikealla puolella hyvin rakennettu koe, joka sulkee pois P:n muut suorat syyt, ja näin ollen rikkoo näennäisen kausaalisen suhteen P:n ja M:n välillä

Interventioilla on lisäksi tiettyjä ominaisuuksia, joita on syytä tarkastella lähemmin. Ensinnäkin interventiot koskevat aina jotain muuttujaa X, ja interventiot määritellään aina suhteessa johonkin toiseen muuttujaan Y. Ei ole olemassa pelkkiä ”tyhjiä” interventioita, vaan aina jonkin X:n interventioita suhteessa johonkin Y:hyn. Tämä johtaa suhteellisuuteen siten, että tietty X:n interventio voi täyttää ehdot suhteessa Y:hyn, mutta ei välttämättä Z:aan. Toiseksi interventiot eivät vaadi inhimillistä toimintaa: intervention ehdot on määritelty ainoastaan kausaalisten suhteiden ja riippumattomuuden käsitteiden kautta. Inhimillinen toiminta voi usein täyttää nämä ehdot, mutta sitä ei vaadita niiden toteuttamiseksi, eikä inhimillisessä toiminnassa ole teorian kannalta mitään erityistä. (Woodward, 2003, ss. 103–104.)

Mikäli muuttuja X on binäärinen, siihen kohdistuvat interventiot ovat helposti ymmärrettäviä: vaihdamme muuttujan arvon joko noltaan tai yhteen, ja katsomme miten jokin muuttuja Y reagoi tähän muutokseen. Kysymys muuttuu vaikeammaksi, jos muuttuja X ei ole näin tarkasti määritelty; interventiota tehdessä on olennaista tietää, mitä muutoksella oikeasti tarkoitetaan. (Woodward, 2003, ss. 112–113.) Esimerkiksi: henkilö X ei tule raskaaksi, koska hän on mies. Tarkastellaksemme tätä väitettä, meidän olisi tehtävä interventio, jossa muutamme X:n ominaisuutta ”olla mies”, ja tarkastelemme tämän vaikutusta hänen mahdollisuuteensa tulla raskaaksi. Ongelma on, että ”mies” ei ole intervention kannalta tarkasti määritelty termi: mitä me tarkalleen ottaen muutamme, jos manipuloimme tätä ominaisuutta? Sukupuoli ei ole binäärinen muuttuja, ja vaikka se olisi, meillä ei ole koherenttia ajatusta siitä, mitä jonkin mieshenkilön muuttaminen naiseksi tosiasiallisesti tarkoittaa. Väite voidaan muotoilla paremmin niin, että se sisältää tarkasti raskaaksi tulemisen mahdollisuuteen vaikuttavat tekijät (esimerkiksi, että henkilöllä on kohtu, yms.) koska näiden yksittäisten asioiden manipulaatiot ovat paremmin määriteltyjä.

Se, että interventiot on määritelty toiminnan sijaan kausaalisten suhteiden kautta sisältää ilmeisen vaaran kehäisyydestä: interventio on itsessään kausaalinen ilmaisu, joten sen avulla ei voida antaa ei-kausaalista selitystä kausaalisille tapahtumille. Onko kausaatio hyväksyttävää selittää käyttämällä apuna toisenlaista kausaatiota? Woodwardin (2003, ss. 104–105) mukaan on, sillä intervention kehäisyys ei ole erityisen ongelmallista tyyppiä.

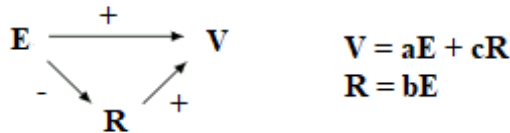
Interventioiden avulla on tarkoitus kuvata kausaalista suhdetta joidenkin X:n ja Y:n välillä, ja interventio itsessään ei ole mitään tämän suhteen olemassaolosta: sen ainoa vaatimus koskee kausaalista suhdetta I:n ja X:n välillä. Woodwardin (2003, s. 105) mielestä käytäntö ja intuitio tukevat ajatusta intervention kehäisyyden vaarattomuudesta: useissa tilanteissa voimme löytää kausaalisen suhteen X:n ja Y:n välillä X:ään kohdistuvat intervention avulla, eikä tässä toimintatavassa ole mitään epäilyttävää tai ongelmallista. Tämä metodi kuitenkin vaatii jonkinlaisen määritelmän siitä, miten X:ää tulee manipuloida siten, että emme oleta mitään X:n ja Y:n välisestä suhteesta ennen kokeen suorittamista. Tähän tehtävään interventiot sopivat erinomaisesti. Lopulta kausaalisen suhteen määrittämisen kannalta olennaista on siis tieto muista tapahtumaan vaikuttavista kausaalisista suhteista ja kontrafaktuaaleista, jotka koskevat Y:n käyttäytymistä, kun X:ään kohdistuu interventio. Kausaatiota ei ole välttämätöntä (tai välttämättä edes mahdollista) redusoida ei-kausaalisiin faktoihin, jotta voimme antaa kausaatiolle ei-triviaalin määritelmän – juuri tähän manipulaatioteoria pyrkii.

2.5 Syiden monimuotoisuus

Olennaisten käsitteiden ollessa paikallaan voidaan siirtyä nyt arvioimaan aiemmin esitettyjä ehtoja kausaalisen suhteen olemassaololle. Riittävätkö **RE** ja **VE** määrittelemään kaikki kausaaliset suhteet tyhjentävästi? Woodward (2003, s. 49) näkee **RE**:n erittäin vakuuttavana: mikäli X:ää manipuloimalla on mahdollista vaikuttaa Y:n arvoon, voidaan sanoa X:n olevan Y:n syy riippumatta siitä, täyttääkö suhde muita tyypillisesti kausaaliselta suhteelta vaadittuja ehtoja, kuten aika-avaruudellista jatkuvuutta tai energian siirtymistä X:n ja Y:n välillä. Jos tämä ehto toteutuu, voimme sanoa kausaalisen suhteen olevan olemassa, mutta voidaanko päätelmä tehdä myös toiseen suuntaan – päteekö **VE** siten, että kaikki syyt täyttävät tämän ehdon?

Tätä kysymystä tarkastellakseen Woodward (2003, ss. 49–50) esittää seuraavan esimerkin. Kuvitellaan, että ehkäisytabletit lisäävät veritulpan todennäköisyyttä. Ne kuitenkin samalla tekevät raskauden hyvin epätodennäköiseksi, ja raskaus puolestaan kasvattaisi merkittävästi veritulpan todennäköisyyttä. Sattumalta kokonaisvaikutus ehkäisytablettien syömisen ja veritulpan todennäköisyyden välillä on 0. Tässä esimerkissä

ehkäisy pillerien suora vaikutus veritulpan todennäköisyyteen kumoutuu täysin sen epäsuorasta vaikutuksesta raskauden todennäköisyyden kautta. Näin ollen ei ole mahdollista toteuttaa yhtä interventiota ehkäisy pillereitä koskevaan muuttujaan siten, että se vaikuttaisi veritulppaa koskevan muuttujan todennäköisyysjakaumaan. Mikäli **VE** määrittäisi hyväksyttävien syiden joukon, tällaisia tilanteita ei voitaisi nähdä hyväksyttävästi kausaalisina. Intuitiomme kuitenkin tuntuu vievän eri suuntaan: tapahtumien välillä vallitsee kausaalinen suhde. **VE** tavoittaa yhden syyn tyyppin, mutta kaikki syyt eivät ole määriteltävissä samalla tavalla: tyhjentävän esityksen on mentävä **RE**:tä ja **VE**:tä pidemmälle.



Esimerkin kausaalinen rakenne ehkäisy pillerien (E), raskauden (R) ja veritulpan todennäköisyyden (V) välillä. Funktionaalisessa esityksessä $a = -bc$, joten E:n suora vaikutus V:hen ja E:n vaikutus V:hen R:n kautta kumoavat toisensa.

VE riittää määrittämään kausaalisen suhteen tilanteissa, joissa syyn kokonaisvaikutus ei ole 0. Tällaista syyn tyyppiä voidaan kutsua kokonaissyiksi (*total cause*) (**KS**), jonka riittävän ja välttämättömän ehdon Woodward määrittelee seuraavasti:

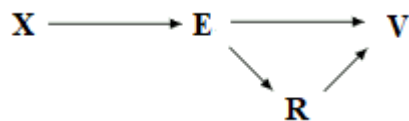
”(**KS**) X on Y:n kokonaissy, jos ja vain jos on olemassa mahdollinen interventio, joka muuttaa X:n arvoa siten, että Y:n arvo (tai sen todennäköisyysjakauma) muuttuu (Woodward, 2003, s. 51)”.

Ehkäisy pillerien syöminen ei siis ole veritulpan kokonaissy. Kausaalinen intuitiomme johtuu siitä, että mikäli raskautta kuvaava muuttuja kontrolloitaisiin muuttumattomaksi (kuten esimerkiksi henkilöiden kohdalla, jotka eivät voi tulla raskaaksi), voisimme tehdä intervention ja vaikuttaa veritulpan todennäköisyyteen. Tämä ajatus tuli esille lyhyesti jo aiemmin, kun annoimme intuitiivisen määritelmän suoralle syyille. Suoran syyn (*direct cause*) (**SS**) käsite määritellään tarkemmin seuraavasti:

(SS) X on Y:n suora syy suhteessa johonkin muuttujien joukkoon V, jos ja vain jos on olemassa mahdollinen interventio, joka muuttaa X:n arvoa siten, että Y:n arvo (tai sen todennäköisyysjakauma) muuttuu, kun kaikki muut V:n sisältämät muuttujat kontrolloidaan interventioilla tiettyihin arvoihin (Woodward, 2003, s. 55).

Kokonaissyyn käsite on monilta osin riittämätön käytännön tieteellisen testaamisen kannalta. Suoran syyn käsitettä tarvitaan kertomaan meille, mitä tekijöitä meidän tulee kokeissa kontrolloida, jotta saamme selville yksin X:n kausaalisen vaikutuksen Y:hyn. Koska myös interventiot ymmärretään vastaavan ajatuksen kautta, suoran syyn käsite on tarpeellinen myös aitojen interventioiden tavoittamiseen. Kokonaissyyt voivat lisäksi vääristää kuvaamme tapahtuman aidosta kausalisesta rakenteesta: ehkäisy pillerien ja veritulpan tapauksessa E:n vaikutus V:hen voidaan helposti määrittää yhtenä mekanismina, joka antaa oikean kausaalisen vaikutuksen, mutta väärällä tavalla – tosiasiasa on olennaista huomioida, että tapahtumaan vaikuttaa yhtäaikaaisesti kaksi erillistä kausaalista ketjua. Kokonaissyyn käsitteellä ei pystytä erottamaan tällaisia tilanteita toisistaan. (Woodward, 2003, s. 52.)

Suorat syyt ja kokonaissyyt kattavat laajan skaalan tapauksia, mutta on olemassa vielä yksi syyn tyyppi, joka ei mahdu näidenkään määritelmien sisään. Kuvitellaan, että edellisessä ehkäisy pillereitä koskevassa esimerkissä olisi olemassa jokin muuttuja X, joka on E:n suora syy. Nyt X ei ole V:n suora syy, sillä mikäli E kontrolloidaan muuttumattomaksi suoran syyn määritelmän mukaisesti, X:n manipulaatiolla ei ole vaikutusta V:n todennäköisyyteen. X ei myöskään aiemmin mainituista syistä ole V:n kokonaissyyn. Tästä huolimatta X tuntuu jollain relevantilla tavalla olevan syy V:n lopulliselle todennäköisyydelle.



X ei ole V:n suora syy, mutta X on V:n osittainen syy

Tässä tapauksessa X voidaan nähdä V:n osittaisena syynä (*contributing cause*) (OS). Osittaiset syyt voidaan määritellä tarkemmin käyttämällä hyväksi aiempaa suoran syyn määritelmää.

(OS) X on Y:n osittainen syy suhteessa johonkin muuttujien joukkoon V, jos ja vain jos (i) on olemassa suorien kausaalisten suhteiden reitti X:stä Y:hyn, ja (ii) on olemassa jokin mahdollinen interventio, joka muuttaa X:n arvoa siten, että Y:n arvo (tai sen todennäköisyysjakauma) muuttuu, kun kaikki V:n muuttujat, jotka *eivät ole* tällä reitillä kontrolloidaan interventioilla tiettyihin arvoihin (Woodward, 2003, s. 59).

Määritelmän mukaisesti esimerkin X voidaan nähdä V:n osittaisena syynä, sillä X:n ja V:n välillä vallitsee suora kausaalinen reitti E:n kautta siten, että mikäli R kontrolloidaan muuttumattomaksi (koska se sijaitsee muuttujajoukossa, mutta ei valitulla reitillä), X:n manipulaatiot vaikuttavat myös V:hen. Huomionarvoista on lisäksi, että kaikki suorat syyt ovat osittaisia syitä, mutta ei päinvastoin.

Osittaisen syyn (ja näin ollen myös suoran syyn) käsite ei kuitenkaan ole täysin ongelmaton. Kuten määritelmästä voidaan huomata, käsitettä voidaan käyttää ainoastaan suhteessa johonkin muuttujien joukkoon V – tämän joukon valinnasta riippuen X voi olla Y:n suora syy joukossa V*, mutta ei välttämättä joukossa V**. Kaikissa tapauksissa tämä ei Woodwardin mukaan ole huono ominaisuus, sillä se vain kuvastaa sitä, millä tasolla haluamme analyysia tehdä. Jos henkilö A ampuu henkilöä B, A:n liipaisimen painallus on B:n kuoleman suora syy. Sama tilanne voidaan kuitenkin kuvata myös tarkemmin määritellyssä joukossa, jossa mukaan otetaan myös aseiden mekaaninen toiminta: liipaisin vapauttaa jousen, joka aiheuttaa iskurin lyönnin, jonka seurauksena patruuna räjähtää ja luoti lähtee lentoon. Tässä tapauksessa liipaisimen painallus ei ole B:n kuoleman suora syy. Muuttujien valinnan jälkeen suoran syyn määrittäminen on kuitenkin objektiivinen kysymys. (Woodward, 2003, ss. 55–56.) Kysymys muuttujien valinnasta voi kuitenkin osoittautua paljon vaikeammaksi.

2.6 Muuttujien valinta ja vakavasti otettavat mahdollisuudet

Muuttujien valintaan liittyviä ongelmia voidaan jälleen tarkastella ehkäisykäsityksien ja veritulpan todennäköisyyttä koskevan esimerkin avulla. Kun tarkasteltava muuttujajoukko sisältää muuttujan R, E on V:n suora syy. Sama ei kuitenkaan päde, mikäli R puuttuu joukosta: E:n vaikutus V:hen on 0, johtuen huomioimatta jääneen muuttujan R vaikutuksesta (Woodward, 2003, s. 56). Muuttujien valinta vaikuttaa merkittävästi siihen, näemmekö jonkin asian syynä vai emme, joten miten tämä valinta tulisi tehdä objektiivisesti? Woodwardin mukaan valinnan täytyy riippua siitä, mitkä tapahtumat näemme kontrafaktuaalisesti vakavasti otettavina vaihtoehtoina olemassa olevalle joukolle, ja että tällainen suhteellisuus on (ja sen kuuluukin olla) kaikkien kausaaliselitysten mallinnusten ominaisuus.

Vakavasti otettavia mahdollisuuksia voidaan kuvata useilla helposti rakennettavilla esimerkeillä. Woodwardin esimerkissä tietty lääkäri on vastuussa potilaasta, jolla on helposti diagnosoitava, antibiooteilla parannettava tauti. Mikäli lääkäri tekee työnsä ja määrää potilaalle antibiootteja, lääkärin toiminta on potilaan parantumisen syy. Vastaavasti, mikäli lääkäri huolimattomuuttaan jättää antibiootit määräämättä, potilaan kuolema on hänen syytään. Kuvitellaan nyt tilanne, jossa lääkäri on toiminnallaan aiheuttamassa potilaan kuoleman. Kaukana sairaalasta asuu henkilö X, joka ei tunne potilasta; tästä huolimatta X voisi päätyä sairaalaan ja huomata, että potilas on jäämässä ilman hoitoa. Hän voisi nyt tehdä oikean diagnoosin ja pelastaa potilaan hengen. Mikäli X jättää tämän tekemättä, onko potilaan kuolema X:n syytä? Tietenkään ei – kaiken tämän tapahtuminen ei ole vakavasti otettava vaihtoehto. Lääkäri puolestaan olisi helposti voinut toimia toisin, ja määrätä potilaan tarvitsemat lääkkeet. Tämä vaihtoehto on vakavasti otettava, ja sen laiminlyöminen tekee lääkäristä vastuullisen. (Woodward, 2003, s. 88.)

Monet asiat vaikuttavat siihen, mitä mahdollisuuksia voimme pitää vakavasti otettavina. Tällaisia asioita ovat ainakin tapahtuman todennäköisyys, moraaliset velvoitteet ja odotukset (lääkärillä on velvollisuus hoitaa potilasta, X:llä ei), miten helposti lopputulos on muutettavissa, tai miten kallista se on. Tällaiset huomiot tuovat välttämättä jonkin verran subjektiivisuutta arvioihin, sillä henkilöiden tarkoitusperät ja toiveet voivat

vaikuttaa siihen, mitkä tapahtumat nähdään vakavasti otettavina. Toisaalta arvioissa on myös objektiivinen puoli: on fakta, että tuntemattomat ihmiset harvoin ilmestyvät sairaalaan pelastamaan potilaita. (Woodward, 2003, s. 89.) Woodward (2003, s. 90) huomioi lisäksi, että suhteellisuuden tuoma subjektiivisuus on vain osittaista, sillä se vaikuttaa ainoastaan meidän tekemiimme kausaalisiin päätelmiin, eikä tarkastelemiemme kontrafaktuaalien totuuksiin. Riippumatta siitä, näemmekö henkilön X saapumisen vakavasti otettavana vai emme, sen johtopäätös on kuitenkin tosi: X olisi voinut pelastaa potilaan. Samoin objektiivisesti tosia tai epätosia ovat muutkin tapahtumaan liittyvät kontrafaktuaalit, olivat ne sitten vakavasti otettavia tai eivät. Subjektiivista on ainoastaan se, mitkä tapahtumat ja kontrafaktuaalit näemme vakavasti otettavina siten, että ne otetaan huomioon kausaalista päätelmää tehdessä. Tämä subjektiivisuus ei kuitenkaan ole äärimmäisen vakavaa, sillä monet valinnat voidaan tehdä joidenkin objektiivisten perusteiden (kuten todennäköisyyksien) pohjalta. Lisäksi se on tarpeellista, sillä meillä on oltava jokin keino sulkea pois tietyt kontrafaktuaalit siten, että esimerkin potilaan kuolemaa ei nähdä X:n syynä. Kaikkien näiden vaihtoehtojen sulkeminen pois tarkastelusta täysin objektiivisesti on mahdotonta, ja ”vakavasti otettavuus” tarjoaa helpon ja intuitiivisen keinon tähän.

Kuinka meidän tulisi tämän perusteella suhtautua ehkäisy pilleriesimerkin muuttujien valintaan? Woodward (2003, ss. 56–57) myöntää, että kysymys jää auki: tällaisissa tilanteissa on puhtaasti kyse muuttujajoukon valitsemisesta, joka jää subjektiiviseksi arvioksi siitä, onko raskautta kuvaava muuttuja vakavasti otettava vai ei. Huomionarvoista kuitenkin on, että molemmat mallit, sisälsivät ne R:n tai ei, ovat yhtä mieltä yksittäisen P:n intervention vaikutuksesta V:hen, ja molemmat ovat tässä suhteessa oikeassa. Tässä tapauksessa satumme tietenkin tietämään, että koska R vaikuttaa merkittävästi V:hen, se tulee nähdä vakavasti otettavana vaihtoehtona ja sisällyttää malliin ainakin henkilöiden kohdalla, jotka voivat tulla raskaaksi; R:n pois jättäminen tässä tilanteessa on virhe. Esimerkki kuitenkin osoittaa, että mikäli mallistamme on jätetty jokin ratkaiseva muuttuja pois, se saattaa vääristää kausaalista kuvaamme tapahtumasta. Tämä on kuitenkin usein tieteellisen toiminnan kohtaama todellisuus: suuri osa tutkijoiden työstä liittyy nimenomaan oikeiden (vakavasti otettavien) muuttujien valintaan. Manipulaatioteorian pyrkiessä olemaan lähellä konkreettista tieteellistä

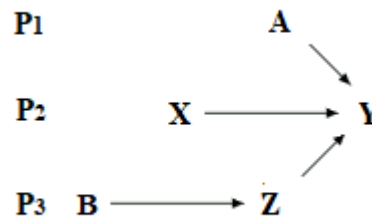
toimintaa, tämän tieteen aspektin sisällyttämistä teoriaan ei voi nähdä ainakaan suurena heikkoutena.

2.7 Manipulaatioteoria

Woodwardin teorian suuri saavutus liittyy kausaalisten merkitysten selventämiseen – pelkkä puhe syistä ja seurauksista jää syiden monimuotoisuuden takia valitettavan epäselväksi. Jotta voidaan antaa tarkka kausaalinen kuvaus tapahtumalle, tarvitsemme tietoa kausaalista rakenteesta, joka puolestaan synnyttää erilaisia kausaalisia suhteita muuttujien välille. Näiden suhteiden kuvaamiseksi tarvitsemme useampaa syyn tyyppiä. **KS** riittää itsessään määrittelemään riittävät ja välttämättömät ehdot kokonaissyiden tapauksessa, ja nyt voimme suoran syyn käsitteen avulla luoda samanlaiset ehdot myös osittaisille syyille (joiden joukkoon myös suorat syyt kuuluvat). Nämä ehdot määrittävät Woodwardin manipulaatioteorian (**M**):

”(**M**) Välttämätön ja riittävä ehto sille, että X on Y :n suora syy suhteessa johonkin muuttujajoukkoon V , on, että on olemassa mahdollinen interventio, joka muuttaa X :n arvoa siten, että Y :n arvo (tai sen todennäköisyysjakauma) muuttuu, kun kaikki muut V :n muuttujat Z_i kontrolloidaan muuttumattomiksi. Välttämätön ja riittävä ehto sille, että X on Y :n osittainen syy suhteessa johonkin muuttujajoukkoon V , on, että (i) on olemassa suorien kausaalisten suhteiden reitti X :stä Y :hyn, eli on olemassa joukko (Z_1, \dots, Z_n) siten, että X on Z_1 :n suora syy, joka puolestaan on Z_2 :n suora syy, joka puolestaan on \dots Z_n :n suora syy, joka puolestaan on Y :n suora syy, ja (ii) on olemassa jokin mahdollinen interventio, joka muuttaa X :n arvoa siten, että Y :n arvo (tai sen todennäköisyysjakauma) muuttuu, kun kaikki muut V :n muuttujat, jotka eivät ole tällä reitillä, kontrolloidaan muuttumattomiksi. Jos on olemassa vain yksi reitti P X :stä Y :hyn, tai jos ainoa vaihtoehtoinen reitti X :stä Y :hyn P :n ohella on suora, niin X on Y :n osittainen syy, kunhan on jokin mahdollinen interventio, joka muuttaa X :n arvoa siten, että Y :n arvo (tai sen todennäköisyysjakauma) muuttuu jollain tietyllä V :n muiden muuttujien arvoilla. (Woodward, 2003, s. 59.)”

Monimutkaisesta muotoilustaan huolimatta **M** pohjautuu hyvin intuitiiviselle ajatukselle: **X** on **Y**:n osittainen syy, mikäli **X**:n muutos aiheuttaa muutoksen **Y**:ssä, kun oikeat muut muuttujat on kontrolloitu tiettyihin arvoihin. Oikeat muut muuttujat määritellään suorien syiden kautta: meidän on kontrolloitava kaikki **Y**:n suorat syyt, jotka eivät löydy tarkasteltavalta reitiltä. (Woodward, 2003, s. 60.) Tämä on lopulta vain tarkempi tapa ilmaista ne vaatimukset, jotka jokaisen hyvin rakennetun kokeen kuuluu käytännön tieteellisessä toiminnassa täyttää: jotta saamme selville yhden muuttujan vaikutuksen, meidän on kontrolloitava kaikki muut vaikuttavat muuttujat *ceteris paribus*-ehdon mukaisesti.



*Mikäli haluamme selvittää, onko **X** **Y**:n osittainen syy, meidän on kontrolloitava **A** ja **Z** muuttumattomiksi, sillä ne ovat **Y**:n suoria syitä, eivätkä ne ole reitillä **P2**. Vastaavasti voimme selvittää, onko **B** **Y**:n osittainen syy kontrolloimalla **X** ja **A**, koska ne eivät ole reitillä **P3**.*

M pyrkii antamaan kuvauksen kausaatiosta manipulaatioiden ja interventioiden kautta. Kausaaliset väitteet tulee ymmärtää interventioita koskevinä väitteinä: ne kertovat, mitä tapahtuisi mikäli asioita muutettaisiin interventioilla – **M** kertoo mitä tapahtuu, mikäli **X** on **Y**:n osittainen syy ja **X**:ää manipuloidaan tiettyjen ehtojen vallitessa. Samalla se antaa riittävät ja välttämättömät ehdot kausaalisen suhteen olemassaololle näiden kahden välillä. **M** kattaa tapaukset, joissa **X** on **Y**:n osittainen syy, ja vastaavasti **KS** kattaa tapaukset, joissa **X** on **Y**:n kokonaissyy. (Woodward, 2003, ss. 60–61.)

Manipulaatioteorian käsitteistön omaksuminen auttaa meitä tulkitsemaan kausaalisia väitteitä, mutta se myös antaa meille keinon arvioida, voidaanko kausaalisia väitteitä tehdä. Filosofinen keskustelu on täynnä esimerkkejä tilanteista, joissa syyn arviointi ei ole yksiselitteistä: kuvitellaan, että salamurhaaja **A** myrkyttää kahvin, jonka hänen uhrinsa

juo ja kuolee. Jos A ei olisi myrkyttänyt kahvia, sen olisi tehnyt salamurhaaja B, ja uhri olisi kuollut joka tapauksessa. Tämä on esimerkki ennalta määräytymisestä (*preemption*), ja se herättää kysymyksen, oliko A:n toiminta uhrin kuoleman syy? Vastaava tilanne syntyy niin kutsutun ylimääräytymisen (*overdetermination*) tapauksessa: nyt A ja B laittavat myrkyt kahviin yhtäaikaaisesti – kumman syy uhrin kuolema nyt on?⁵ (Hitchcock, 2007.) Manipulaatioteoria antaa yhden tavan analysoida näitä tilanteita, ja auttaa selkiyttämään harhailevaa intuitiotamme: molemmissa tapauksissa voidaan soveltaa suoran syyn käsitettä, ja sen avulla voidaan hahmottaa, mitkä ovat tapahtuman kannalta relevantit kontrafaktuaalit, joiden totuutta arvioimalla voimme tehdä kausaalisen päätelmämme – lähtökohtaisesti, että ensimmäisessä tapauksessa A:n toiminta on uhrin kuoleman syy, ja että toisessa tapauksessa molempien toiminta on uhrin kuoleman syy.⁶

Varustettuna käsityksellä siitä, mitä kausaatiolla ja sen eri tyypeillä tarkoitetaan, siirrytään seuraavaksi tarkastelemaan miten näitä määritelmiä voidaan käyttää selitysten tuottamiseen.

3 Kausaaliselitykset

3.1 Historiallinen katsaus ja ei-kausaaliset selitysteoriat

Woodwardin teoriassa selitykset ovat linkittyneitä manipulaatioon ja tapahtumien kontrollointiin liittyviin kysymyksiin: jonkin tapahtuman selitys yksinkertaisesti kertoo meille, miten tähän tapahtumaan voidaan vaikuttaa. Historiallisesti filosofiset selitysteoriat ovat keskittyneet manipulaatiosuhteita enemmän yhdistämään selitykset ennusteisiin, odotettavuuteen ja tapahtumien systematisaatioon. Woodwardin selitysteoria on erityisesti sosiaalitieteiden kannalta valtava kehitysaskel, ja jotta sen

⁵ Nämä esimerkit eivät ole tyypikkauseksaalisia väitteitä, joten niihin täytyy soveltaa yksittäisiin tapauksiin soveltuvaa manipulaatioteorian muotoilua (ks. Woodward, 2003, p. 84) – esimerkit on kuitenkin verrattain helposti ymmärrettävissä myös esitetyssä tyypitason kehityksessä. Olennainen ero on, että emme voi kontrolloida muita muuttujia täysin vapaasti, sillä yksittäisen tapauksen piirteet on pääpiirteittäin säilytettävä.

⁶ Relevantti kontrafaktuaali ensimmäisessä tapauksessa on se, että toista myrkytystä ei tapahdu, jolloin kuolema on riippuvainen A:n toiminnasta. Toisessa tapauksessa puolestaan se, että mikäli toinen myrkytyksistä ei tapahdu (oli se kumpi tahansa), kuolema on riippuvainen myrkyttäjän toiminnasta. Kausaalisen päätelmän tekeminen ei silti ole täysin yksiselitteistä, ja sitä ohjaa subjektiivinen näkemysmekanismemme vakavasti otettavuudesta alaluvun 2.6 käsittelyn mukaisesti.

ratkaisevat erot ja edut saadaan nostettua esiin, on syytä tarkastella sitä taustaa, jota vasten Woodward oman teoriansa kehitti.

3.1.1 DN/IS-mallit

Hempelin deduktiivis-nomologinen (DN) malli on yksi tieteellisen selittämisen teoretisoinnin keskeisimmistä saavutuksista. Hempelin esiin nostamat huomiot ja erityisesti hänen mallinsa kohtaama kritiikki on ohjannut lähes kaikkea seurannutta keskustelua, ja se luo korvaamattoman viitekehyksen, johon myöhempiä malleja voidaan peilata.

Hempelin mukaan tapahtuma voidaan selittää näyttämällä, miksi sen tapahtuminen oli odotettavissa tilanteeseen liittyvien relevanttien lakien ja alkutilan perusteella. DN-mallissa selitettävä asia (*explanandum*) selitetään deduktiivisena argumenttina, jonka premisseinä (*explanans*) on joukko alkutilaa kuvaavia deskriptiivisiä väitteitä ja vähintään yksi relevantti laki, joka kertoo miten nämä väitteet suhteutuvat toisiinsa (Hempel, 1965, s. 336). Esimerkiksi auringon näennäisen pieni koko on odotettavissa, kun otetaan huomioon deskriptiiviset seikat auringon koosta ja auringon ja maan välimatkasta, sekä optinen laki välimatkan vaikutuksesta havaittuun kokoon. Voimme siis deduktiivisena argumenttina esittää, miksi aurinko näyttää meille pieneltä.

$$\begin{array}{rcl}
 & L_1, \dots, L_k & \text{Lait } (k > 0) \\
 & C_1, \dots, C_m & \text{Alkuehdot} \\
 \text{Deduktio} & \frac{\quad}{E} & = \text{Explanans} \\
 & & \text{Selitettävä tapahtuma} = \text{Explanandum}
 \end{array}$$

DN- selitys – selitys muodostuu koko argumentista (Niiniluoto, 1983, s. 274)

Arjessa selitykset eivät useinkaan täytä näitä deduktiivisen argumentin kriteerejä: kun selitämme lapselle, miksi aurinko näyttää pieneltä, toteamme vain, että aurinko on todella kaukana. Mainitsematta jää explanansin muut osat, ja joudumme olettamaan, että lapsi osaa itse täyttää selityksen puuttuvat palat. Tällaiset epätäydelliset selitykset ovat *selitysluonnoksia*, joiden selityksellinen voima kuitenkin perustuu niiden taustalla olevaan aitoon deduktiiviseen selitykseen (Woodward, 2017). Vastaavia tilanteita nousee

esiin erityisesti, kun kaikkia selitettävään tapahtumaan vaikuttavia tekijöitä ja lakeja on mahdotonta ilmaista edes niin halutessamme.

Lakien sisältyminen päättelyyn on DN-mallissa niin keskeistä, että se tunnetaan myös peittävän lain mallina. Tosiasioita voi selittää ja ymmärtää ainoastaan, kun on tiedossa laki, jonka sovellusalaan tapahtumat kuuluvat. Lakien keskeisyys on kuitenkin myös heikkous, sillä mallin pitäisi vastata vakuuttavasti myös kysymykseen siitä, mikä on selityksen kannalta hyväksyttävä laki. Tämä kysymys vaatii vastausta siihen, miten luonnonlait määritellään, ja vastaus tähän on osoittautunut äärimmäisen vaikeasti tavoitettavaksi. Valtava määrä filosofista kirjallisuutta on syntynyt aiheen ympäriltä keskittyen usein X:n määritelmään yhtälössä: luonnonlaki = yleinen totuus + X (yleisesityksenä tästä lähestymistavasta ja sen ongelmista ks. Dretske, 2004). X:n tehtävä on poistaa sattuman mahdollisuus lain määritelmästä. On olemassa lukematon määrä tapauksia, jotka ovat yleisesti tosia vain sattumalta, ja joita voidaan näennäisesti käyttää lakeina DN-selityksissä – kuten esimerkiksi Hempelin (1965, s. 339) oma toteamus, että kaikki Greenburyn koulun johtokunnan jäsenet vuonna 1964 ovat kaljuja. Tämä ”laki” ei varmasti voi selittää yhdenkään tietyn jäsenen kaljuutta hyväksyttävästi.

Mikä sitten erottaa aidot lait näennäisistä laeista ja sattumalta tosista yleistyksistä? Voimme nostaa esiin ainakin muutaman ominaisuuden, joita aidoilla luonnonlaeilla tulee olla. Ensinnäkin lain on oltava yleinen siinä mielessä, että se pätee universaalisti ilman yksilöitä koskevia poikkeuksia. Toiseksi lain tulee tukea kontrafaktuaalista tarkastelua: sen tulee kyetä vastaamaan ”mitä jos tilanne olisi ollut erilainen”-kysymyksiin. Tämän lisäksi lailla on oltava modaalista vaikutusta: sen tulee rajata sitä, mikä on mahdollista tai mahdotonta. (Salmon, 1989, s. 14.) Etenkin modaalinen vaikutus ja tuki kontrafaktuaaleille tavoittavat jotain olennaista luonnonlaeista. Jos koulun johtokuntaa koskevalla lakiväitteellä olisi nämä ominaisuudet, se tarkoittaisi, että uusi jäsen muuttuisi välittömästi ja välttämättä liittyessään kaljuksi modaalisen vaikutuksen takia, ja voisimme myös arvioida kontrafaktuaalisesti mikä vaikutus johtokuntaan liittymisellä tai sieltä poistumisella on henkilön hiustyylisiin. Ongelma on, että nämä ominaisuudet muodostavat niin tiiviin kokonaisuuden, että pelkästään niiden avulla on mahdotonta arvioida, onko mikään yksittäinen väite aito luonnonlaki. Väitteellä on modaalista vaikutusta, ja se tukee

kontrafaktuaalista tarkastelua vain, jos se on luonnonlaki; vastaavasti väite on luonnonlaki vain, jos sillä on modaalista vaikutusta, ja se tukee kontrafaktuaalista tarkastelua.

Hempel tiesi itsekkin, että DN-malli vaatii selostuksen siitä, mitä hyväksyttävällä lailla todellisuudessa tarkoitetaan. Filosofit ovat vuosikymmeniä pyrkineet luomaan metafysisesti johdonmukaista ja episteemisesti käyttökelpoista teoriaa laeista, mutta tyhjentävää vastausta ei ole vielä olemassa. Emme tässä pyri löytämään tarkkaa lainomaisuuden määritelmää, joten tässä vaiheessa tyydytään toteamaan, että määritelmän puute on vakava ongelma DN-mallille.

DN-mallia voidaan soveltaa tilanteisiin, jossa lait takaavat sisältämiensä tapahtumien varmuuden, mutta epävarmoihin tilastollisiin lakeihin perustuvat selitykset tuottavat sille vakavia ongelmia. Hempelin (1965, ss. 381–382) klassinen esimerkki koskee John Jonesia, joka kärsii bakteerin aiheuttamasta kurkkukivusta. Hän ottaa tautiinsa penisilliiniä ja toipuu nopeasti. DN-malli on nyt kyvytön selittämään penisilliinin ja Jonesin paranemisen yhteyttä: ei ole olemassa yleistä lakia, jonka mukaan kaikki, jotka kärsivät samasta bakteerista ja ottavat penisilliiniä, toipuvat nopeasti.

Vastauksena ongelmaan syntyi IS-malli (*inductive-statistical*), joka pohjautuu monilta osin samanlaisille periaatteille kuin DN-mallikin: selitys muodostuu argumentista, joka kertoo, miksi explanandum oli odotettavissa explanansin perusteella. Koska IS-mallissa käytettävä laki on tilastollinen, odotettavuus ei kuitenkaan ole enää deduktiivista varmuutta, vaan induktiivista todennäköisyyttä. Kuinka IS-selitys Jonesin parantumiselle voisi rakentua? Oletetaan, että on olemassa seuraava tilastollinen yleistys: jokaisella, joka kärsii bakteerin aiheuttamasta kurkkukivusta, ja joka ottaa penisilliiniä, on korkea parantumisen todennäköisyys. Otamme tämän lakiväitteeksi, ja lisäämme alkuehdot, joiden mukaan Jones kärsii bakteerin aiheuttamasta kurkkukivusta ja ottaa penisilliiniä. Nyt voimme selittää parantumisen induktiivisena argumenttina, jonka mukaan Jonesin parantuminen oli odotettavaa (eli todennäköistä) lakiväitteen ja alkutilan perusteella.

$$\text{Induktio } \frac{\text{P}(G | F) = r}{\frac{\text{Fa}}{\text{Ga}}} [r]$$

IS-selitys: 1. premissin lakiväitteen mukaan G:n todennäköisyys, kun F pätee, on r.

2. premissin alkuehdon mukaan F pätee kohteelle a.

Täten voimme päätellä, että G pätee kohteelle a todennäköisyydellä r. (Salmon, 1989, s. 53.)

Koska IS-malli perustuu ajatukselle johtopäätöksen odotettavuudesta, tapahtuman todennäköisyyden on oltava lähellä yhtä. Mitä korkeampi tapahtuman todennäköisyys on, sitä suuremman induktiivisen tuen se argumentille antaa. Toisin kuin DN-selitys, IS-selitys ei takaa johtopäätöksen totuutta: minkään yksittäisen tapahtuman ei ole mukauduttava johtopäätökseen siitä huolimatta, että lain mukaan johtopäätös on odotettavissa.

3.1.2 Puuttuva kausaatio

Kuten jo johdannossa todettiin, syyn ja seurauksen käsitteet ovat olennaisia ainakin arjen selityksen kannalta. Hempelin selitysmallit eivät kuitenkaan eksplisiittisesti mainitse kausaatiota, vaan DN/IS-mallien on tarkoitus kattaa laajempi joukko selityksiä, joista kausaaliselitykset muodostavan vain osan (von Wright, 1973, s. 15). Tässä näkemyksessä kausaaliselitykset ovat vain DN-selityksiä, joissa käytetään kausaalilakeja. DN-mallia vastaan esitetyt kuuluisat vasta-argumentit kuitenkin osoittavat, että tämän seurauksena malli ei sisäistä kausaation keskeistä roolia lähes kaikkien selitysten kannalta.

Ensinnäkin, mallit mahdollistavat selityksiä, jotka eivät vaikuta aidoilta selityksiltä. Tunnetun vasta-argumentin mukaan DN-mallissa lipputangon pituus voidaan selittää sen varjon pituuden kautta, vaikka intuitiivisesti selityksellinen suhde liikkuu toiseen suuntaan – lipputanko luo varjon, ei toisinpäin (Bromberger, 1966, ss. 92–93). Yleiset lait ja funktionaaliset suhteet voidaan kuitenkin muotoilla tällä tavoin symmetrisesti, jolloin malli tuottaa selityksiä, jotka eivät ole aidosti hyväksyttäviä.

Toiseksi, osa hyväksyttävältä tuntuista selityksistä ei täytä DN-mallin vaatimuksia. Kuvitellaan seuraava tilanne: potkaisen vahingossa pöytää, minkä seurauksena sen päällä oleva mustepullo tipahtaa, ja sen sisältämä muste leviää matolle. Matolla olevalle mustetahralle tuntuu nyt löytyvän hyvin ilmeinen selitys, joka koostuu tahraan johtaneiden tapahtumien kuvauksesta. Selitys tuntuu aidolta, vaikka se ei noudata DN-mallia laisinkaan: emme pystyisi edes halutessamme kertomaan kaikkia tilanteeseen vaikuttaneita relevantteja lakeja ja päättämään seurausta deduktiivisesti. (Scriven, 1962, ss. 198–199.) Voimme tietenkin vedota aiemmin mainittuun ajatukseen selitysluonnoksesta, mutta tämä nostaa vain esiin erilaisia ongelmia. Esimerkin selitysluonnos tuntuu lisäävän ymmärrystämme, mutta miten tämä on mahdollista, vaikka sen takana piilevä aito selitys on jo määritelmällisesti saavuttamattomissa? Selityksellä on arvoa ymmärrystä tuottavana toimintana, ja tämän mukaisesti ainoastaan aidot selitykset voivat luoda ymmärrystä. Ei siis tunnu perustellulta, että pelkästään jonkin piilevän DN-mallisen selityksen olemassaolo riittää tämän ymmärryksen tuottamiseen, vaikka emme saisi sitä koskaan selville. (Scriven, 1962, ss. 205–208.)

Kolmantena, DN-malli tuntuu hyväksyvän selityksiä, jotka ovat täysin irrelevantteja itse tapahtuman kannalta. Kuvitellaan lakiväite, jonka mukaan kukaan ehkäisy pillereitä syövä mies ei tule raskaaksi, ja alkutila, jonka mukaan John Jones syö ehkäisy pillereitä. Näin ollen voimme päätellä, että Jones ei tule raskaaksi. Tämä on hyväksyttävä DN-selitys, jonka päättelyn mukaan on odotettavaa, että Jones ei tule raskaaksi, koska hän syö ehkäisy pillereitä. (Salmon, 1989, s. 50.) Selitys on kuitenkin jälleen epäkelpo puuttuvan kausaation takia: selittävä tekijä ei tosiasiallisesti aiheuta selitettävää asiaa millään tavalla.

Premissien relevanssiin liittyvät kysymykset aiheuttavat myös toisenlaisia ongelmia etenkin IS-mallin kohdalla. Uusi informaatio voi muuttaa induktiivisen päätelmän johtopäätöstä, joten mikäli lisäämme aiempaan esimerkkiin premissin, jonka mukaan Jonesin bakteeri on resistentti penisilliinille, johtopäätöksemme muuttuu täysin päinvastaiseksi. Ongelmaksi muodostuu referenssiluokan valinta: Jones saattaa kuulua joukkoon, joka kärsii normaalista bakteerista, tai hän saattaa kuulua joukkoon, joka kärsii resistentistä bakteerista. Tilanne ja todennäköisyydet muuttuisivat vielä lisää, jos saisimme tietää, että Jones on hyvin vanha, tai että hänellä on hyvin terveet elämäntavat.

Voimme luoda mahdollisia referenssiryhmiä lähes loputtomasti, joten kysymys kuuluu, mikä niistä meidän tulisi valita? Voimme pyrkiä tekemään valinnan mahdollisimman rationaalisesti, mutta lopulta kaikki IS-selitykset ovat tämän valinnan myötä episteemisesti relativistisia, ja niiden totuus riippuu ennalta määrätystä taustatiedosta, joka kertoo mitkä tekijät ovat relevantteja päättelymme kannalta. Tämä tarkoittaa myös, ettei IS-malli voi koskaan taata sisältämänsä päättelyn totuutta: sillä ei yksin ole keinoja erottaa relevantteja tekijöitä maailmasta.

Vaikka DN/IS-malleista on nykykeskustelussa luovuttu lähes kokonaan, niiden panos tieteellisen selittämisen teorialle on ollut mittaamattoman arvokas. Tarjoamalla helpon kritiikin kohteen ne ovat tehneet aiheesta helposti lähestyttävän, ja luoneet lähtökohdat niitä seuranneiden teorioiden rakentamiselle. Tarkastellaan seuraavaksi kahta kenties tärkeintä yritystä korjata DN/IS-mallien ongelmia, joiden sisältämiä ajatuksia Woodward päätyi omassa teoriassaan tarkentamaan.

3.1.3 Unifikaatio

Vaikka suuri osa DN/IS-mallien ongelmista liittyy puuttuvaan kausaatioon, kaikki filosofit eivät lähteneet rakentamaan teorioitaan kausaation roolin kasvattamisen kautta. Erityisesti Philip Kitcher pyrki luomaan teoriaa, joka pohjautuu kausaation sijaan unifikaatioon. Kitcherin (1989, s. 430) mukaan hyväksyttävä selitys on sellainen, joka kuuluu tiettyyn selitysten luokkaan, joka muodostaa selityksellisen varastomme (*explanatory store*). Selityksellinen varasto muodostuu tieteen väitteistä, jotka yhdessä luovat parhaan systematisaation olemassa olevasta tiedostamme. Täten selitysteorian tärkein tehtävä on määrittää, mitkä selitykset kuuluvat selitykselliseen varastoon, ja avain tähän parhaaseen systematisaatioon löytyy unifikaatiosta.

Karkeasti Kitcherin ajatus on seuraava: selityksessä on tarkoitus johtaa kuvauksia mahdollisimman monesta eri tapahtumasta käyttämällä mahdollisimman vähän erilaisia argumenttimuotoja. Paradigmaattinen esimerkki on maanpäällisten ja taivaankappaleita koskevien liiketeorioiden yhdistäminen: unifioimalla nämä teoriat, tarvitsemme vähemmän argumenttimuotoja antamaan kuvaukset suuremmasta määrästä tapahtumia. (Kitcher, 1989, ss. 434–435.) Tiede kehittyy unifikaation kautta, ja sen myötä

perimmäisten faktojen määrä maailmasta pienenee. Selityksellinen varasto puolestaan muodostuu niistä argumenttimuodoista, jotka sillä hetkellä maksimaalisesti unifioivat sen hetken tieteelliset uskomukset.

Kitcherin teoria ei sisällä kausaatiota, joten miten se pystyy vastaamaan DN-mallia vastaan esitettyihin vasta-argumentteihin? Kitcher (1989, s. 485) itse käsittelee lipputankoesimerkkiä seuraavasti. Kuvitellaan, että meidän selityksellisessä varastossamme on argumenttimuoto, jonka avulla voimme päätellä asioiden pituuksia riippuen niiden syntytavasta ja kokemista muutoksista. Tämän muodon mukaan voimme päätellä lipputangon pituuden sen rakentajan aikomuksista ja käyttämistä materiaaleista, ja päädyimme alustavaan johtopäätökseen lipputangon pituudesta sen syntyhetkellä. Kun otamme huomioon lipputangon alkuperäisen pituuden, ja kaikki sen tämän jälkeen kokemat muutokset, saavumme johtopäätökseen lipputangon nykyisestä pituudesta. Lipputangon varjon pituuden voimme selittää samalla tavoin viittaamalla varjon syntyhetkeen. Halutessamme päätellä lipputangon pituuden sen varjon kautta meidän on nyt lisättävä selitysvarastoomme toinen argumenttimuoto, jonka mukaan voimme päätellä asioiden pituuksia niiden varjojen perusteella. Aiempi selityksellinen varastomme pystyy kuitenkin jo selittämään kaikkien olioiden pituuden, joten tämän uuden muodon lisääminen ei kasvata selitettävien tapahtumien määrää – se vain kasvattaa käytettävien argumenttimuotojen määrää. Mikäli luopuisimme aiemmasta syntyävän ja muutoksen argumenttimuodosta varjomuodon hyväksi, selitettävien tapahtumien määrä pieneneisi, koska kaikki asiat eivät luo varjoja. Täten varjomuoto ei kuulu selitykselliseen varastoomme eikä se ole hyväksyttävä selitys. Huomioitavaa on toki, että varjomuoto tarjoaa joka tapauksessa käyttökelpoisen tavan päätellä asioiden pituuksia – Kitcherin unifikaation kannalta tämä päättely ei kuitenkaan tuota *selityksiä*.

Kitcherin näkemyksessä selitykset johdetaan mahdollisimman unifioiduista teorioista. Mikään unifikaatioteoriassa ei vaadi, että kaikki selitykset ovat deduktiivisia, mutta samalla se tuntuu hyvin helposti johtavan tällaiseen näkemykseen. Kausaation rooli teoriassa on puolestaan täysin alisteinen selitykselle: kausaaliset huomiomme pohjautuvat täysin unifikaatioon perustuville selityksellisille suhteille. (Woodward, 2003, s. 360.) Kausaation jättäminen teorian ulkopuolelle johtaa kuitenkin jälleen ongelmiin kausaalisen

intuitiomme takia. Kitcherin lipputankoesimerkin käsittelyssä keskeinen oletus on, että kaikki asiat eivät luo varjoja. Mikäli kaikilla asioilla olisi varjo, voisimme hyvin vaihtaa syntyvän ja muutoksen argumenttimuodon varjomuotoon, sillä selitettävien tapahtumien määrä ei tällöin vähenisi, ja argumenttimuotojen määrä pysyisi samana. Intuitiomme kuitenkin sanoo, että vaikka kaikki asiat loisivat varjoja, joiden perusteella voisimme saada selville kaikki niiden ulottuvuudet, olisi silti epäsovpevaa selittää asioiden pituus niiden varjojen kautta. (Woodward, 2003, s. 361.)

Unifikaatio tavoittaa jotain keskeistä aidoista selityksistä, mutta puuttuvan kausaation takia se ei pysty sisäistämään ajatusta, että tosiasiallisesti kausaatio – *ei* unifikaatio – on selityksen ratkaisevin osuus. Toisin sanoen: unifikaatio on hyvien kausaalisten selitysten ominaisuus, ei päinvastoin, ja kausaalisen teorian on pyrittävä unifiamaan tieteellisiä selityksiämme.

3.1.4 SR-malli

Wesley Salmon pyrki SR-mallissaan (*statistical relevance*) luomaan tilastollisen selityksen teorian, joka tavoittaa kausaatiosta sen, mihin IS-malli ei pystynyt. Kuvitellaan nyt, että Jones kärsii hyvin sitkeästä bakteerista, josta paranee itsestään vain prosentin todennäköisyydellä, ja penisilliinin ottaminenkin nostaa paranemisen todennäköisyyden vain kymmeneen prosenttiin. IS-malli on kykenemätön selittämään penisilliinin ja paranemisen yhteyttä, vaikka Jones penisilliiniä otettuaan sattuisi paranemaan: lopputulos ei ole yksiselitteisesti odotettavissa. Mikäli kuitenkin otamme lähtökohdaksemme korkean todennäköisyyden sijaan tilastollisen relevanssin, penisilliinillä voidaan nähdä olevan selvä vaikutus paranemiseen ja paraneminen voidaan penisilliiniä ottaneiden kohdalla selittää tämän kautta.

SR-malli pyrkii erottamaan aidot selitykset tarkastelemalla, miten mahdollinen selitys vaikuttaa tapahtuman todennäköisyyteen. Formaalisti ilmaistuna vertaamme tapahtuman priorista todennäköisyyttä sen posterioriseen todennäköisyyteen. Priorisessa todennäköisyydessä tarkastellaan selitettävän tapahtuman itsenäistä todennäköisyyttä, ja posteriorisessa selitettävän tapahtuman ehdollista todennäköisyyttä, kun mahdollinen selittävä asiantila on tapahtunut. Mikäli tapahtuman posteriorinen todennäköisyys ei ole

sama kuin sen priorinen todennäköisyys, selitys on hyväksyttävä. (Psillos, 2002, ss. 252–253.) Eli jos X on selitettävä tapahtuma, ja E on mahdollinen selitys tapahtumalle, niin jos $P(X|E) \neq P(X)$, niin E on hyväksyttävä selitys tapahtumalle X (Salmon, 1989, s. 67). Tästä syystä penisilliini hyväksytään selitykseksi Jonesin paranemiselle, sillä muutoksen ei tarvitse olla iso, jotta ehto täyttyy.

Kuinka sitten määritämme, mikä todennäköisyys valitaan prioriseksi ja mikä posterioriseksi? Salmonin (1989, s. 63) mukaan priorinen todennäköisyys mitataan ennen referenssiluokan ositusta ja posteriorinen sen jälkeen. Jonesin tapauksessa hän kuuluu ensin referenssiryhmään, joka kärsii sitkeästä bakteerista, joten priorinen todennäköisyys muodostuu itsenäisen paranemisen todennäköisyydestä. Tämä referenssiluokka voidaan lisäksi jakaa kahteen: niihin, jotka saavat penisilliiniä, ja niihin, jotka eivät. Posteriorinen todennäköisyys muodostuu siis paranemisen todennäköisyydestä penisilliinin ottamisen jälkeen. Voimme kuitenkin jatkaa tätä ajatusta pidemmälle, sillä myös penisilliiniä ottaneiden ryhmä voidaan jakaa kahteen: esimerkiksi nuoriin ja vanhoihin. Etsiessämme selitystä, meidän on kuljettava tämä ketju loppuun asti, ja sijoitettava Jones oikeaan referenssiluokkaan siten, että sitä ei enää pysty relevantisti osittamaan. Tällaista luokkaa kutsutaan homogeeniseksi. SR-selitys muodostuu koko ketjun perusteella: selitys on joukko tapahtuman kannalta tilastollisesti relevantteja faktoja. Yksittäinen relevantti seikka voidaan nähdä eräänlaisena osaselityksenä tapahtumalle.

SR-malli pyrkii tavoittamaan kausaation sen kautta, että tilastollinen relevanssi implikoi usein kausaalista suhdetta. SR-malli ei kuitenkaan sisällä mitään viittausta kausaatioon, vaan se rakentuu pelkästään tilastollisten yhteyksien ja korrelaatioiden varaan. Tämä tuottaa ongelmia tilanteissa, joissa kaksi muuttujaa ovat korreloituneita, mutta johtuvat yhteisestä syystä. Tätä voidaan kuvata jo aiemmin kohdatulla esimerkillä ilmapuntarista ja myrskystä: kun ilmapuntari näyttää matalia lukemia, myrskyn todennäköisyys kasvaa (Psillos, 2002, s. 254). Tosiasiassa ilmapuntarin lukema on irrelevantti myrskyn tapahtumisen kannalta, sillä sekä lukeman että myrskyn aiheuttaa lasku ilmanpaineessa. Ilmanpaineen sanotaan näin ollen varjostavan (*screen off*) ilmapuntarin ja myrskyn suhteen, ja SR-selityksiä etsiessä on vaadittava, ettei mikään ulkoinen tekijä varjosta tarkasteltavien muuttujien suhdetta.

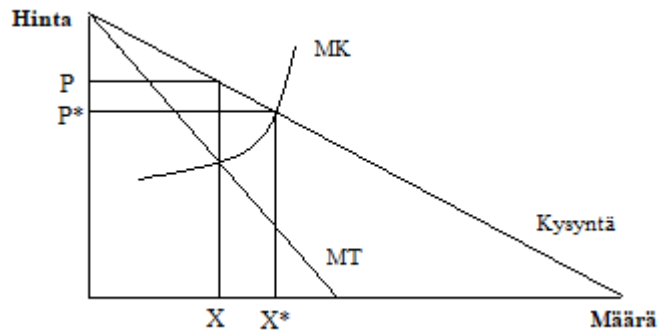
SR-malli on selvä kehitysaskel IS-mallista, mutta se myös jää askeleen päähän niistä tavoitteista, joiden varaan se alun perin rakentui. Ensinnäkin, SR-malli kärsii samoista referenssiluokan valintaan liittyvistä ongelmista kuin IS-malli. Emme koskaan voi olla täysin varmoja siitä, olemmeko löytäneet homogeenisen referenssiryhmän, ja tämä valinta on aina suhteellista meidän olemassa olevaan tietoomme mahdollisesti relevanteista referenssiryhmän osituksista. (Salmon, 1989, s. 77.) Lisäksi kausaalisuhteiden ja tilastollisten relevanssien yhteys ei ole niin yksiselitteinen kuin mallin kannalta voisi toivoa. Tilastollinen relevanssi on usein kausaalisten suhteiden alimääräämä: erilaiset kausaalisuhteet voivat tuottaa samanlaisia relevanssisuhteita. Esimerkiksi tilanne, jossa ilmapuntarin lukema aiheuttaa matalapaineen, joka puolestaan aiheuttaa myrskyn, luo täysin samanlaiset relevanssisuhteet kuin tilanne, jossa matalapaine on ilmapuntarin lukeman ja myrskyn yhteinen syy. SR-mallin ei toimiakseen tarvitse kertoa, mitkä kausaalisuhteet ovat voimassa. (Woodward, 2017.) Täten se epäonnistuu tavoitteessaan, mikäli sen tärkein motivaatio koskee selitettävän asian kannalta merkittävien kausaalisten tekijöiden löytämistä.

Voimme lisäksi todeta, että kuten unifikaation kohdalla, SR-malli kääntää tilanteen väärin päin: tilastolliset relevanssit ovat arvokkaita vain, koska ne kertovat meille niiden pohjalla olevista kausaalisuhteista. Tosiasiassa mallin selitysvoima perustuu kausaaliin suhteisiin, eikä suinkaan pelkkään tilastolliseen relevanssiin. SR-malli luo tilastollisen pohjan selitykselle, mutta sitä on vahvistettava kausaalisilla tekijöillä, jotta voimme tuottaa hyväksyttäviä tieteellisiä selityksiä. Toisin sanoen: tilastollinen relevanssi tulee selittää kausaalisten suhteiden kautta, ei toisin päin. (Salmon, 1989, s. 166.)

Päätämme tähän historiallisen katsauksen ja palaamme tarkastelemaan James Woodwardin kausaaliselityksen manipulaatioteoriaa. Woodwardin teoria pystyy säilyttämään suuren osan aiemmin esitettyjen mallien vahvuuksista vastaten samalla niiden kohtaamiin ongelmiin. Sen avulla voimme myös luoda erinomaisen viitekehyksen tieteelliselle selittämislle myös sosiaalitieteiden saralla.

3.2 DN-mallista manipulaatioihin

Aloitetaan Woodwardin teorian tarkastelu palaamalla lyhyesti DN-mallin pariin. Kuvitellaan seuraava tilanne: näemme ikkunasta korpin, ja haluamme antaa selityksen sen mustalle värille. Tiedämme alkutilan, jonka mukaan lintu on korppi, ja satumme tietämään myös, että kaikki korpit ovat mustia (oletetaan esimerkin vuoksi, että tämä on hyväksyttävä laki). DN-mallia seuraten voimme argumenttina esittää perustelun linnun mustalle värille, ja olemme samalla antaneet sille selityksen. Kuvitellaan nyt toinen tilanne, jossa pohdimme miksi monopolit johtavat hyvinvointitappioon verrattuna täydellisen kilpailun markkinoihin. Kun monopoli ottaa haltuunsa jonkin tuotannonalan, se kohtaa koko markkinoiden kysynnän, ja voi valita, miten se haluaa tähän kysyntään vastata. Tästä seuraa, että hinnat nousevat ja tuotettu määrä laskee. Tämä tyypillinen mikrotaloustieteen esimerkki voidaan esittää seuraavasti.



Käyrä MT kuvaa marginaalituottoa, eli tuottoa, jonka monopoli saa, kun se tuottaa yhden yksikön lisää, ja käyrä MK kuvaa marginaalikustannusta, eli kustannuksen nousua yhden lisäyksikön tuottamisesta. Täydellisen kilpailun tilanteessa tuottajat eivät pysty itse vaikuttamaan tuotettuun kokonaismäärään tai hintaan, joten tasapaino löytyy tilanteesta, jossa MK ja kysyntä kohtaavat, eli tuotetta tuotetaan määrällisesti X^* hinnalla P^* . Niin kauan, kun tuottajilla on mahdollista saavuttaa voittoja tuottamalla lisää, näin tehdään, kunnes voitot painuvat nolliin. Monopolitilanteessa tätä ei tapahdu: monopoli voi itse määrittää sekä tuotteen hinnan että valmistetun määrän, eli se saa itse päättää, missä pisteessä se kohtaa kysynnän. Tämä piste löytyy kohdasta, jossa MT ja MK risteävät. Tämän pisteen oikealla puolella yhden lisäyksikön tuottaminen maksaa enemmän kuin

mitä se tuottaa, ja sen vasemmalla puolella lisäyksikön tuottaminen kasvattaa voittoja. Voitot maksimoiva tasapaino löytyy siis pisteestä (X, P), jossa tuotetaan vähemmän kuin kilpailullisessa tilanteessa, ja kuluttajien maksama hinta on korkeampi. (Woodward, 2003, ss. 189–190.)

Tämänkaltainen selitys vaikuttaa hyväksyttävältä ja lisäksi syvällisemmältä kuin edeltävä linnun väriä koskeva tapaus. Se ei kuitenkaan täytä DN-selityksen vaatimuksia. Vaikka monopolitilanteen selitys voidaan kuvata argumenttina, jonka mukaan hyvinvointitappio on odotettavissa, sen käyttämä lakiväite ei voi edes teoriassa täyttää aidon lain vaatimuksia – selityksen pohjalla oleva yleistys, jonka mukaan tuottajat maksimoivat voittoa, ei täytä yhtäkään lakien vaatimuksista. Linnun väriä koskeva esimerkki puolestaan täyttää DN-mallin vaatimukset, mutta vaikuttaa tästä huolimatta epätydyttävältä selitykseltä, sillä se ei tunnu tuottavan ymmärrystä – palaamme tähän ajatukseen tarkemmin seuraavassa alaluvussa 3.3.

Woodward (2003, s. 191) näkee ratkaisevan eron näiden selitysten välillä olevan niiden kyky vastata kontrafaktuaalisiin kysymyksiin siitä, miten selitettävä tapahtuma olisi muuttunut, mikäli alkutilanne olisi ollut erilainen. Toisin sanoen, hyvät selitykset mahdollistavat lopputuloksen vaihtoehtojen tarkastelun ja kertovat, miten selitettävä tapahtuma riippuu johdonmukaisesti selityksessä kuvatuista tekijöistä. Voittojen maksimointia koskevan oletuksen ei tarvitse olla yleinen laki, jotta voimme sen avulla kertoa, mitä tapahtuisi, jos esimerkiksi monopolin tuotannon kustannukset nousisivat (tuotettava määrä laskisi ja hinta nousisi entisestään), tai mikäli ala vapautettaisiin kilpailulle. Antaessamme selitystä sille, miksi monopolilla on tällainen vaikutus tuotettuun määrään, olennaista on kertoa, miten tilanne muuttuu monopolin ottaessa markkinat haltuunsa. Korppiesimerkissä selitykseltä puuttuu tämä ominaisuus: sen avulla ei voida kertoa, missä tilanteessa korppi olisi jonkin muun värinen kuin musta, ja sen avulla ei voida päätellä, kuinka korpin väriä voisi muuttaa (Woodward, 2003, s. 193). Se ei toisin sanoen kerro linnun värin kausaalisista syistä. Yleisesti lintujen väriyksen voidaan analyysin tasosta riippuen nähdä johtuvan evolutionaarisista tai biologisista syistä, ja tällaiset esimerkiksi genetiikkaan pohjautuvat selitykset kertovat myös, kuinka

väriin voisi vaikuttaa. Nämä selitykset täyttävät manipulaatioteorian vaatimukset, ja ne myös vaikuttavat intuitiivisesti aiempaa esimerkkiä tyydyttävimmiltä.

DN-selitykset ovat täten Woodwardin mallissa vakavasti puutteellisia – ei riitä, että selitys kertoo, miksi lopputulos on nomologisesti odotettavissa. Selitysten avulla on lisäksi pystyttävä vastaamaan erilaisiin ”mitä jos tilanne olisi ollut erilainen?”-kysymyksiin (jatkossa *m-kysymykset*), ja tämä on mahdollista kuvaamalla selitettävän tapahtuman syitä siten, niin kuin ne manipulaatioteorian mukaisesti ymmärretään (Woodward, 2003, s. 195). Tämä ajatus on intuitiivisesti hyväksyttävissä, mutta tarkastellaan seuraavaksi lyhyesti, miksi juuri *m-kysymykset* ovat niin relevantteja selitysten (ja ymmärryksen tuottamisen) kannalta.

3.3 Selityksellinen voima

Kun sanomme manipulaatioteorian selitysten olevan parempia kuin DN-mallin selitykset, missä suhteessa tämä paremmuus mitataan? Korpin värin selittäminen edeltävällä tavalla vaikuttaa epätydyttävältä, koska se ei tunnu auttavan meitä ymmärtämään, miksi korppi tosiasiallisesti on musta. Tämä tarkoittaa, että jollain relevantilla tavalla ymmärryksen muodostuminen on riippuvaista siitä, miten selitys onnistuu kuvaamaan lopputuloksen ja alkutilan välistä riippuvuussuhdetta. Tämä ei kuitenkaan ole ainoa huomionarvoinen seikka arvioidessamme selitysten kykyä tuottaa ymmärrystä, eli niiden selityksellistä voimaa.

Kuorikoski ja Ylikoski (2010) tunnistavat viisi ulottuvuutta, joiden mukaan selityksellistä voimaa voidaan arvioida. Ensimmäinen näistä on epäherkkyys (*non-sensitivity*): selitys on sitä voimakkaampi, mitä kestävämpi se on suhteessa muutokseen alkutilassa. Epäherkkyys vaikuttaa olevan se ulottuvuus, joka kaikkein eniten erottaa DN- ja manipulaatiomallien selitykset toisistaan. Kuvaamalla selitettävään tapahtumaan johtavia kausaalisuhteita, manipulaatiomallin selitykset ovat tyypillisesti hyvin epäherkkiä suhteessa muutokseen alkutilassa: vastaamalla *m-kysymyksiin* selitykset kattavat myös tilanteet, joissa alkutila muuttuu, ja kertovat miten lopputulos reagoi tähän. DN-malli puolestaan pystyy selittämään vain yhden alkutilan, johon sen sisältämä lakiväite on sovellettavissa. Sen avulla ei voida vastata *m-kysymyksiin*, joten DN-mallin selitykset

ovat äärimmäisen herkkiä, ja suoriutuvat erityisen heikosti tämän keskeisen ulottuvuuden kohdalla.

Muut Ylikosken ja Kuorikosken kuvaamat ulottuvuudet ovat: explanandumin tarkkuus, faktuaalinen tarkkuus, integraation aste olemassa olevaan tietoon ja kognitiivinen helppous. Explanandumin tarkkuus (*precision*) viittaa siihen, miten hyvin selitettävä tilanne on määritelty, eli mihin skaalaan suhteutettuna selitys annetaan. Esimerkiksi, ”miksi henkilö valitsi tummansinisen auton muiden värien sijaan”, on tarkemmin määritelty explanandum kuin ”miksi henkilö valitsi sinisen auton muiden värien sijaan”. Vastaavasti, ”miksi henkilö valitsi sinisen auton muiden värien sijaan” kattaa laajemman skaalan kontrasteja, kuin ”miksi henkilö valitsi sinisen auton punaisen auton sijaan”. Mitä laajempi skaala kontrasteja, ja mitä tarkemmin määritelty explanandum, sitä parempi selitys. (Ylikoski & Kuorikoski, 2010, ss. 210–211.)

Faktuaalinen tarkkuus selityksellisenä ulottuvuutena kuvaa sitä, miten todenmukainen selitys on. Tämän arviointi on olennaista esimerkiksi silloin, kun selitykset pohjautuvat yksinkertaistettuihin malleihin. Tähän paneudutaan syvemmin luvussa 4 taloustieteellisten selitysten yhteydessä. Integraation aste olemassa olevaan tietoon puolestaan tavoittaa tärkeimmän osan unifikaatioteorian selityksellisestä voimasta: selitykset, jotka kuuluvat osaksi laajempaa kokonaisuutta ja ovat helposti yhdistettävissä siihen, ovat yleensä vahvempia kuin osiensa summa. (Ylikoski & Kuorikoski, 2010, ss. 212–213.)

Viimeinen ulottuvuus, kognitiivinen helppous, kuvaa sitä, miten helposti selitys on omaksuttavissa: miten helppo sen päättelyä on seurata, miten helppo sen sovellusala on ymmärtää, ja miten helposti sen vahvuudet ja heikkoudet on käsitettävissä (Ylikoski & Kuorikoski, 2010, ss. 214–215). DN-malli pärjää varsin hyvin tällä ulottuvuudella: linnun värin selitys ”korppiuden” kautta on helppo tapa selittää väri nopeasti ja yksinkertaisesti. Tämä helppous kuitenkin syntyy muiden ulottuvuuksien kustannuksella. Tämä havainto kuitenkin auttaa ymmärtämään, miksi DN-mallin tuottamilla selityksillä on eräänlaista arvoa tietyissä tarkkaan rajatuissa tilanteissa, vaikka ne tuntuvatkin yleisesti epätydyttäviltä muiden ulottuvuuksien vajavaisuuden takia. Manipulaatioteorian

mukaisen näkemyksen omaksumisessa ei ole kyse ainoastaan kausaation sisällyttämisestä teoriaan, vaan myös siitä, että tämän myötä voimme luoda objektiivisesti voimakkaampia selityksiä.

3.4 Selitys ja kontrafaktuaalit

Vastauksemme m-kysymyksiin ovat tosiasiasa kuvauksia tietyistä kontrafaktuaaleista ja arvioita niiden totuusarvoista. Selityksen kannalta relevantit kontrafaktuaalit ovat niitä, jotka kuvaavat interventioiden seurauksia. (Woodward, 2003, s. 196.) Monopoliesimerkissä tämä tarkoittaa intervention ehtojen täyttäviä manipulaatioita, joilla vaikutetaan esimerkiksi kysyntään, kustannuksiin, tai firman monopoliasemaan.

Selitysten yhdistäminen konkreettisiin manipulaatioihin ja niiden seurauksiin ratkaisee kausaalisen asymmetrian aiheuttamat ongelmat, joihin DN-malli ei pysty antamaan tyydyttävää vastausta. Vaikka voimme millä tahansa hetkellä kertoa lipputangon pituuden funktionaalisesti sen varjon pituuden avulla, lipputangon pituus ei tosiasiasa ole kontrafaktuaalisesti riippuvainen varjon pituudesta. Tämän selvittäminen on konkreettisesti mahdollista ja vieläpä hyvin helppoa – varjon pituuden muuttaminen ei vaikuta lipputangon pituuteen. Lipputankoon kohdistuva interventio puolestaan muuttaa varjon pituutta. Tämän avulla voimme vastata skaalaan m-kysymyksiä siitä, mitä varjon pituudelle tapahtuisi, jos lipputanko olisi eri pituinen. Havaitun asymmetrian takia voimme sanoa, että lipputanko selittää varjonsa pituuden, ei toisin päin.

Voimme tarkastella muita DN-mallia kohtaan esitettyjä vastaväitteitä samalla tavalla. Kaatuneen mustepullon tapauksessa voimme selvästi huomata, että mustetahrin syntyminen on kontrafaktuaalisesti riippuvaista jalan osumisesta pöytään. Näin arkisessa tilanteessa ei ole välttämätöntä kuvata kaikkia tapahtumaan vaikuttaneita tekijöitä, vaan selitykseksi riittää yksinkertainen havainto siitä, että *jos* jalkani ei olisi osunut pöytään, mustepullo ei olisi kaatunut. Vastaava tulkinta voidaan antaa myös selityksellisten irrelevanssien kohdalla. John Jonesin ehkäisytablettien syöminen – ja siihen kohdistuvat interventiot – eivät vaikuta hänen raskaaksi tulemisensa todennäköisyyteen lainkaan. Tämä yksinkertainen käsittely toimii kokonaissyiden tarkasteluun, mutta monimutkaisemmissa kausaalirakenteissa selitykset ja irrelevanssi on tulkittava aiemmin

esitetyn manipulaatioteorian määritelmän mukaisesti. Esimerkiksi ehkäisytablettien ja veritulpan tapauksessa tarkasteltavat kontrafaktuaalit ovat niitä, joissa raskauden todennäköisyys on interventiolla kontrolloitu muuttumattomaksi (Woodward, 2003, s. 200).

Lopulta päädyimme DN-mallin kohdalla samanlaiseen päätelmään kuin unifikaatio- ja SR-mallien kohdalla. DN-mallin kuvaama lakeihin perustuva odotettavuus ei ole riittävää aidosti voimakkaiden selitysten kannalta, ja tapahtumaa kuvaavien propositioiden deduktiivinen rakenne on merkityksellinen vain silloin, kun sen avulla voidaan tarkastella kontrafaktuaaleja ja vastata m-kysymyksiin. Manipulaatioteorian mukaisesti fyysiset riippuvuussuhteet ja niiden kuvaus ovat selitysten fundamentaalinen osa, ja deduktiivisilla päätelmillä ei ole tämän kannalta itsenäistä selityksellistä arvoa; ne ovat arvokkaita vain, kun ne kuvaavat oikein niiden taustalla olevia aitoja kausaalisuhteita. (Woodward, 2003, s. 202.) Manipulaatioteoriassa selitykset ovat vain kuvauksia kontrafaktuaalisista riippuvuussuhteista tapahtumien välillä, joita voimme löytää ja tarkastella interventioiden avulla.

3.5 Selityksiä ilman lakeja

Kuten aiempi monopoliesimerkki osoittaa, monet selityksemme perustuvat yleistyksiin, joiden avulla voimme vastata m-kysymyksiin niiden sovellusalan sisällä. Oletus voittojen maksimoinnista ei selvästikään täytä luonnonlain ehtoja, mutta se tuntuu kuitenkin jollain tavalla lain kaltaiselta. Kaikkien selitysten pohjalta ei kuitenkaan löydy edes tällaista lainomaiselta vaikuttavaa väitettä.

Otetaan jälleen yksinkertainen esimerkki. Koulutuksen ja tulotason välinen suhde on yksi taloustieteen tutkituimmista riippuvuussuhteista. Tätä on tyypillisesti mallinnettu Mincer-yhtälön avulla, jossa tuntiansioiden logaritmi ($\ln(y)$) esitetään koulutusvuosien (S) ja työkokemuksen (X) regressiona seuraavasti: $\ln(y) = a + bS + cX + dX^2 + e$. Vakiotermi a kuvaa tulotason, kun koulutus- ja työvuodet ovat nolla. Virhetermi e puolestaan kattaa mahdolliset mallin ulkopuoliset vaikuttajat. Datan avulla voidaan nyt estimoida arvoja regressiokertoimille, jotta saamme selville yksittäisten tekijöiden vaikutuksen, jotka voivat vaihdella ajasta ja paikasta riippuen. Björklund ja Kjellström

(2002) arvioivat parametreille seuraavat arvot käyttäen Ruotsin dataa vuodelta 1991: $a = 3,596$, $b = 0,046$, $c = 0,026$ ja $d = -0,037$. Koska malli on log-lineaarinen, koulutusvuosien vaikutus voidaan tulkita siten, että yhden lisäkoulutusvuoden vaikutus henkilön tuntiansioihin on 4,6 prosenttia.

Yllä kuvattu vaikutus tulotason ja koulutusvuosien välillä tahdotaan nähdä kausaalisenä. Monella tapaa se myös vaikuttaa intuitiivisesti selittävältä, mutta enää selityksen pohjalla olevaa yleistystä ei voida nähdä juuri millään tavalla lainomaisena. Kun lukitsemme parametriarvot datan perusteella saatuihin estimaatteihin, kuvaus muuttujien suhteesta pätee vain yleisellä tasolla, ja vain hyvin tarkasti rajatussa ajassa ja paikassa: tarkalleen ottaen ainoastaan Ruotsissa, ja ainoastaan vuonna 1991. Woodwardin (2003, s. 207) mielestä tämä kuitenkin riittää selitykseen, niin kauan kun tarkastelu todellakin rajataan oikeaan aikaan ja paikkaan, ja esitetty malli kuvaa (riittävän) todenmukaisesti tiettyjä voimassa olleita aitoja kontrafaktuaalisia riippuvuussuhteita. Voimme yllä olevan mallin avulla kertoa, mitä keskiarvoisesti tapahtuu, jos manipuloimme Ruotsissa vuonna 1991 asuvan työssäkäyvän koulutus- tai työvuosien määrää, ja voimme huomata tällä olevan suora vaikutus henkilön keskimääräiseen tulotasoon. Tällaisia yleistyksiä voidaan käyttää rajatusti vastaamaan m-kysymyksiin, ja ne ovat näin ollen myös selittäviä tämän rajauksen sisällä.

Selityksiksi sopivat yleistyksset ovat siis sellaisia, joiden avulla voidaan vastata m-kysymyksiin. Luonnonlait, jotka ovat universaalisti tosia yleistyksiä, ovat tietenkin tällaisia. Woodwardin keskeinen havainto kuitenkin on, että universaalius tai lainomaisuus eivät ole välttämättömiä selityksen kannalta. Selityksen kannalta ratkaisevaa on ainoastaan, että yleistys pysyy voimassa silloinkin, kun sen sisältämien muuttujien arvoja manipuloidaan – eli että sen avulla voidaan vastata m-kysymyksiin. Luonnonlait täyttävät tämän vaatimuksen, mutta niin täyttävät myös lukuisat muut yleistykset, jotka eivät täytä muita lainomaisuuden ehtoja. Selityksen kannalta ratkaisevaa ei olekaan lainomaisuus, vaan *invarianssi*.

3.6 Invarianssi

Invariantti yleistys on sellainen suhde kahden muuttujan välillä, joka pysyy vakaana tiettyjen interventioden tapahtuessa. Vakaudella tarkoitetaan yleistuksen kykyä kuvata vähintään likimääräisen todenmukaisesti sen sisältämien muuttujien suhdetta silloinkin, kun näihin muuttujiin kohdistetaan joitakin interventioita. Tämä ajatus muodostaa kontrastin aiempaan näkemykseen laeista, joiden on pysyttävä vakaana *kaikkien* muutosten kohdalla. Luonnonlait ovat automaattisesti invariantteja, ja historiallisesti lainomaisuuden on nähty edeltävän invarianssia: mikäli yleistys hajoaa minkään muutoksen kohdalla, yleistys ei ole laki eikä täten invariantti lainkaan. Woodwardin (2003, s. 243) mukaan meidän tulisi tämän sijaan nähdä invarianssi suhteellisena käsitteenä, joka muodostuu jostain alarajasta, ja tämän ylittävästä jatkumosta. Mikäli yleistys ei pysy vakaana minkään interventioden kohdalla, se jää rajan alapuolelle: yleistys ei ole invariantti eikä täten käyttökelpoinen selitykseksi. Rajan ylittävät yleistykset puolestaan ovat invariantteja suhteessa johonkin skaalaan mahdollisia muutoksia. Aiempi esimerkki on osoitus näistä invarianssin eri asteista: fyysiset luonnonlait ovat invariantteja ajasta ja paikasta riippumatta, kun taas kuvaus tulojen ja koulutusvuosien suhteesta voi olla invariantti vain suhteessa muutoksiin, jotka kohdistuvat sen sisältämiin muuttujiin Ruotsissa vuonna 1991.

Invarianssin asteet nousevat esiin myös niissä interventioissa, joita voidaan kohdistaa yleistuksen kuvaamiin muuttujiin. Edellisen esimerkin yhtälö (estimoiduilla parametriarvoilla) on invariantti suhteessa sen sisältämien muuttujien interventioihin niin kauan, kun nämä muutokset ovat järkevien rajojen sisällä. Yhtälö kuvaa todenmukaisesti tulotason muutosta, oli koulutusvuosia nolla tai kymmenen, mutta mikäli jollain henkilöllä olisi yli sata koulutusvuotta, hänen kohdallaan yhtälön kuvaama suhde hajoaisi varmasti. Yleistystä voidaan kuitenkin käyttää selitykseen niiden tapausten kohdalla, joiden suhteen se on invariantti, vaikka tämä invarianssi hajoaisi tietyssä pisteessä. Mitä invariantimpi yleistys on, sitä suurempaan määrään m-kysymyksiä sen avulla pystytään vastaamaan – eli sitä epäherkempi se on suhteessa muutoksiin alkutilassa, ja sitä paremman selityksen se tarjoaa (Woodward, 2003, s. 243).

Ideaalissa tapauksessa invarianssin asteiden vertailu on helppoa. Galilein lain avulla voidaan kuvata funktionaalisesti jonkin kappaleen putoamiseen kuluva aikaa, jos tiedetään, miltä korkeudelta kappale pudotetaan. Tämä yleistys on invariantti tiettyjen pudottamiskorkeuteen vaikuttavien interventioiden kohdalla, mutta mikäli kappale pudotetaan äärimmäisen korkealta (suhteessa maapallon säteeseen), tai mikäli maapallon massa olisi merkittävästi erilainen, yleistys lakkaisi toimimasta. Newtonin lait kattavat myös nämä tapaukset: ne ottavat lisäksi huomioon sen kappaleen massan ja säteen, jonka päältä pudotus tapahtuu: ne ovat invariantteja suhteessa laajempaan skaalaan interventioita kuin Galilein lait. Newtonin lait antavat näin ollen syvällisemmän selityksen putoamisajalle kuvaamalla paremmin niitä tekijöitä, joista selitettävä asia todellisuudessa riippuu. Kun vertaamme kahta yleistystä tällä tavalla, teemme sen hyvin tarkasta näkökulmasta, joka liittyy vain invarianssiin suhteessa tiettyihin interventioihin. Galilein ja Newtonin tapauksessa tutkimme vain, miten vakaana yleistykset pysyvät suhteessa interventioihin, jotka muuttavat pudotuskorkeutta tai alla olevan kappaleen sädettä tai massaa. Tämä on aivan eri asia kuin sen tutkiminen, onko jokin teoria suoraan johdettavissa toisesta. (Woodward, 2003, ss. 261–262.)

Harvat yleistykset kuitenkaan muodostavat edeltävän kaltaista tilannetta, jossa toisen invarianssin skaala on täydellinen osajoukko toisesta. Monet yleistykset, joita haluamme vertailla keskenään, eroavat olennaisesti juuri siinä, minkä interventioiden suhteen ne ovat invariantteja. Miten voimme uskottavasti vertailla tällaisia tilanteita toistensa kanssa? Woodwardin (2003, s. 262) mukaan tilanteesta riippuen tietyt invarianssin ulottuvuudet ovat olennaisempia kuin toiset: voimme vertailla yleistyksiä riippuen siitä, miten invariantteja ne ovat suhteessa näihin ulottuvuuksiin, joita pidämme kaikkein tärkeimpinä tapahtuman selityksen kannalta. Yksilöiden toimintaa selitetään usein yksinkertaisella rationaalisen päätöksenteon mallilla, jonka mukaan henkilöt maksimoivat hyvinvointiaan ja agenttien preferenssit ovat transitiivisia. Malli on näin ollen invariantti suhteessa muutoksiin agentin informaation määrässä tai hänen kohtaamisissaan hinnoissa – muutosten tapahtuessa agentti vain mukauttaa tapansa maksimoida hyvinvointiaan. Malli ei kuitenkaan ole invariantti esimerkiksi suhteessa agentin aivokemiaan tehtäviin muutoksiin, joilla hänestä muokataan epäitsekäs. Tämä invarianssin ulottuvuus ei kuitenkaan vaikuta tärkeältä rationaalisen päätöksenteon mallin tavoitteleman selityksen

kannalta. (Woodward, 2003, s. 263.) Invarianssin ulottuvuuksien eriarvoisuus mahdollistaa vertailun yleistysten välillä. Nämä ulottuvuudet luonnollisesti vaihtelevat tutkimusalasta riippuen, ja valittu ulottuvuus eräällä tavalla määrittää selityksellisen agendamme: evoluutiobiologiassa ei pyritä selittämään kuluttajien valintoja, ja vastaavasti taloustieteessä ei pyritä selittämään, miten meistä on kasvanut taloudellisia olentoja (Woodward, 2003, s. 262).

Yleistyksen voidaan toimia selityksinä, jos ne ovat invariantteja relevantilla tavalla, ja eri yleistyksiä voidaan vertailla suhteessa niiden invarianssin luonteeseen ja asteeseen. Se, mitä ”enemmällä” tai ”tärkeämmällä” invarianssilla lopulta tarkoitetaan, jää näitä ajatuksia hyödyntävien erityistieteiden vastuulle.

Edellä on lyhyesti esitetty Woodwardin selitysteorian ydin. Seuraavassa luvussa pyrimme soveltamaan näitä ajatuksia taloustieteelliseen mallinnukseen, mutta ennen tätä on syytä lyhyesti katsoa taaksepäin: millaisia edistysaskeleita manipulaatioteoria on ottanut suhteessa aiempiin teorioihin?

3.7 Edistysaskeleita

Manipulaatioteorian on tarkoitus vastata moniin aiempien mallien kohtaamiin ongelmiin. Kuten olemme huomanneet, suuri osa DN-mallin ongelmista katoaa, kun ymmärrämme selitykset kuvauksina kausaalisista suhteista. Pyritään seuraavaksi lisäksi tarkastelemaan, miten manipulaatioteoria pystyy sisäistämään vahvuuksia sekä unifikaatio- että SR-mallista, ja mahdollisesti täyttämään niiden puutteita. Lopuksi tahdon sen historiallisen merkityksen vuoksi esitellä lyhyesti Salmonin CM-mallin, ja samalla perustella tulevan taloustieteitä koskevan luvun kannalta, miksi manipulaatioteoria soveltuu tämän työn tarkoituksiin myös CM-mallia paremmin.

3.7.1 Unifikaatio

Kitcherin epätydyttävää vastausta lipputankoesimerkkiin käsiteltiin aiemmin, ja kuten DN-mallin tapauksessa, myös tähän verrattuna manipulaatiomalli pystyy tarjoamaan paremman vastauksen. Unifikaatioteoria kärsii kuitenkin myös toisenlaisista vaikeuksista.

Unifikaatio käsitteenä on monitulkintainen ja kattaa useita erilaisia tieteellisiä saavutuksia ja tavoitteita (Woodward, 2003, ss. 362–363). Biologiassa unifikaatio voi tarkoittaa uuden yleisen luokituksen luomista, matemaattisissa tieteissä unifikaatiolla voidaan ymmärtää erilaisten ongelmien tulkitsemista saman viitekehyksen avulla, ja lisäksi unifikaatio voi olla fysikaalista siten, että aiemmin erillisiksi luultujen tapahtumien ymmärretään johtuvan samasta mekanismista. Näistä ainoastaan fyysinen unifikaatio tuntuu liittyvän kausaaliselityksiin, sillä ainoastaan se pyrkii tavoittamaan jotain tapahtumien pohjalla olevista kausaalirakenteista. Tämä herättää Woodwardin (2003, ss. 363–364) mukaan kysymyksen siitä, voiko unifikaatioteoria erotella fyysisen unifikaation selittävät väitteet muun tyyppisen unifikaation deskriptiivisistä väitteistä. Vaikka loisisimme uuden luokittelukeinon, esimerkiksi että X on nisäkäs, jonka avulla voimme sijoittaa suuremman määrän yksilöitä johonkin luokkaan, tämä ei varsinaisesti ole selittävää toimintaa. Vaikka kaikilla nisäkkäillä on sydän, ja voimme nisäkkäiden luokittelun avulla määrittää, että jollain tietyllä olennolla on sydän, koska se on nisäkäs, tämä selitys tuntuu epätyytyttävältä. Kuten aiemmin korpin värin tapauksessa, luokittelun avulla ei voida vastata kysymykseen siitä, mitä tapahtuisi asioiden ollessa toisin.

Toinen ongelma unifikaation kannalta on, että se ei pysty selittämään heikompien selitysten arvoa. Galilein lain avulla pystytään selittämään tippuvien kappaleiden nopeutta, vaikka se on vähemmän unifioitu kuin Newtonin teoria. Tämän tulisi tarkoittaa, että Galilein laki on vain vähemmän selittävä kuin Newtonin teoria, mutta unifikaatioteoriassa ei näytä olevan tilaa vähemmän selittäville teorioille: teoria joko kuuluu selitykselliseen varastoon tai ei. (Woodward, 2003, ss. 367–368.) Selityksellinen varasto luo aiemmin kohtaamamme alarajan selityksille, mutta tämä alaraja näyttää olevan virheellisesti aseteltu. Mikäli luovumme alarajasta, kohtaamme toisenlaisen ongelman: kaikki huonotkin selitykset ovat tietyllä tapaa unifioivia, joten ilman alarajan määrittävää selityksellistä varastoa, unifikaatioteorialla ei ole enää johdonmukaista keinoa määrittää, mitkä selitykset ovat hyväksyttäviä. Woodwardin (2003, ss. 368–369) mukaan unifikaatioteorian kannattajat joutuvat valitsemaan toisen: ”joko ei ole olemassa rajaa (kaikki selitykset ovat hyväksyttäviä, mutta vähemmän unifioidut ovat vähemmän selittäviä), tai ei ole olemassa jatkumoa (vain parhaiten unifioidut selitykset ovat aitoja selityksiä)”.

Unifikaatioteorian vahvuus on se, että sen vaatimus yleisyydestä tuntuu tavoittavan jotain keskeistä aidoista selityksistä. Manipulaatioteoriassa tämä ajatus kiteytyy invarianssin käsitteeseen. Invarianssi antaa vastauksen sekä rajanvetotapaukseen että yleistysten vertailuun alarajan ylittävässä jatkumossa. On kuitenkin painotettava, että invarianssi ei ole vain uusi termi unifikaatiolle: invarianssin asteiden ja unifikaation vertailu ovat kaksi erillistä asiaa. Tästä huolimatta ei liene sattumaa, että hyvin invariantit yleistyksiset ovat usein myös yleisiä ja unifioivia unifikaatioteoriaa vastaavassa mielessä. (Woodward, 2003, s. 368.) Tämä on myös todennäköinen syy unifikaatioteorian kannattajien tekemään attribuutiovirheeseen: useat hyvät selitykset ovat unifioivia, mutta ne eivät ole hyviä selityksiä vain, koska ne ovat unifioivia. Tosiasiassa selityksen voimaa ei arvioida vain integraation asteen ulottuvuudella, ja unifikaatio pärjää huonosti muiden arvioitavien ulottuvuuksien kohdalla (Ylikoski & Kuorikoski, 2010, s. 213).

Woodwardin teorialla on keinot vastata moniin unifikaatioteorian ongelmiin, ja lisäksi sen avulla voidaan huomioida myös yleisyyden tärkeys selitysten kannalta. Yleisyys on olennaista kahdella tavalla: selitys on sitä parempi, ensinnäkin mitä useampaan m-kysymykseen sen avulla voi vastata (epäherkkyys), ja toiseksi mitä paremmin se voidaan yhdistää olemassa olevaan tietoon, ja kuinka tämän yhdistämisen avulla voidaan vastata yhä laajempaan skaalaan erilaisia m-kysymyksiä. Selitys on selitys, koska se täyttää manipulaatioteorian vaatimukset, mutta selitys on tietyllä ulottuvuudella sitä parempi, mitä unifioivampi se on.

3.7.2 SR-malli ja todennäköisyydet

SR-mallin tärkein saavutus koski ehdollisen todennäköisyyden ja selitysten yhteyden paljastamista. Mallista jäi kuitenkin puuttumaan kriittinen askel, sillä kausaaliselitysten näkökulmasta yhteys ehdollisten todennäköisyyksien ja selityksen välillä on vain näennäinen: todennäköisyyksillä on tosiasiassa yhteys kausaalirakenteeseen, ja vasta tämän kautta selityksiin. Tarkastellaan seuraavaksi, kuinka manipulaatioteoria pystyy korjaamaan tämän puutteen.

Klassinen todennäköisyyksiä hyödyntävä kausaation määritelmä löytyy esimerkiksi Cartwrightilta (1983, s. 26): ”E aiheuttaa X:n, jos ja vain jos E kasvattaa X:n

todennäköisyyttä kaikissa tilanteissa, jotka ovat muuten kausaalisesti homogeenisiä suhteessa X:ään.” Vaatimus kausaalisesta homogeenisyydestä tarkoittaa, että muut X:ään vaikuttavat kausaaliset tekijät on kontrolloitava muuttumattomiksi. Woodward (2003, s. 62) huomioi, että tällaisella näkemyksellä on joitakin eroja hänen malliinsa verraten: ensinnäkin, ehto on ilmaistu ehdollisten todennäköisyyksien eikä kontrafaktuaalien kautta, ja toiseksi, se on määritelty koskemaan vain kaksiarvoisia muuttujia. Voidaan huomata myös, että tämä ehto pystyy tavoittamaan ainoastaan positiivisia kausaalisia vaikutuksia. Woodwardin mallissa positiivisen vaikutuksen sijaan tavoitellaan kausaalista relevanssia, joka voi olla myös negatiivista: yhteys on syytuuttujan ja seurausmuuttujan todennäköisyysjakauman välillä. Kuinka Cartwrightin ehtoa voidaan muuttaa siten, että se kattaa laajemman skaalan tapahtumia Woodwardin kuvaamalla tavalla?

Cartwrightin muotoilun on yksinkertaisesti tarkoitus antaa määritelmä sille systemaattiselle yhteydelle, mikä vallitsee kausaalisten suhteiden ja ehdollisten todennäköisyyksien välillä. Mikäli haluamme laajentaa tätä projektia jatkuviin muuttujiin ja kausaaliin relevansseihin, Woodwardin (2003, s. 63) mukaan ratkaisut olisivat muodoltaan seuraavanlaisia: ”E aiheuttaa X:n jos ja vain jos E ja X ovat riippuvaisia toisistaan ehdollisesti suhteessa tiettyihin muihin muuttujiin F”. Palautetaan nyt mieleen SR-mallin ehto selitykselle: E on hyväksyttävä selitys X:lle, mikäli $P(X|E) \neq P(X)$. Tämä ehto vastaa Woodwardin ratkaisua tilanteessa, jossa ei ole muita muuttujia F. Mikäli haluamme tarkastella tilannetta muiden muuttujien F kanssa, SR-mallin ehto muuttuu muotoon $P(X|E, F) \neq P(X|F)$, joka myös vastaa Woodwardin näkemystä. SR-mallin suurin ongelma koskee sitä, ettei se pysty johdonmukaisesti kertomaan, mitkä muuttujat F on otettava huomioon tätä arviota tehdessä. Manipulaatioteoria puolestaan antaa tähän hyvin suoran keinon. Riippuen siitä, näemmekö E:n olevan X:n osittainen vai kokonaissyys, valitsemme kontrolloitavat muuttujat vastaavan määritelmän mukaan. Näiden kontrollien ollessa paikallaan (eli kun X:n todennäköisyysjakauma on ehdollinen muuttujiin F nähden), E on X:n syy, mikäli näiden välillä edelleen vallitsee riippuvuussuhde. (Woodward, 2003, s. 65.) Tätä riippuvuussuhdetta voidaan tarkastella X:n manipulaatioiden avulla, ja se täyttää SR-mallin ehdot hyväksyttävälle selitykselle.

SR-mallissa ehdolliset todennäköisyydet kuvaavat riippuvuussuhteita muuttujien arvojen välillä tiettyjen ehtojen vallitessa. Manipulaatiomalli pystyy sisällyttämään tämän ajatuksen kuvaamalla tapahtuman kausaalista rakennetta, joka ilmenee (tarkasteltavan rakenteen mukaisten kontrollien ollessa paikallaan) SR-mallin tyyppisinä tilastollisina riippuvuussuhteina. Tässä tapauksessa relevanssit ovat kuitenkin vain kausaalisten suhteiden ominaisuus, ja selityksellinen voima perustuu näihin kausaalisiin suhteisiin tilastollisen riippuvuuden sijaan.

3.7.3 CM-malli

CM-malli (*causal mechanical*) syntyi Wesley Salmonin yrityksestä täydentää SR-mallin puutteita antamalla määritelmä kausaatiolle prosessien kautta. Kausaalinen prosessi on Salmonin (1989, s. 108) mukaan jatkuva fyysinen prosessi, jolla on kyky siirtää jokin merkki siten, että tämä muutos pysyy voimassa itse tapahtuman jälkeenkin. Salmon antaa tästä esimerkin: majakka lähettää ympärilleen valkoista valoa jatkuvana fyysisenä prosessina. Mikäli muokkaamme valoa siten, että vaihdamme majakkaan punaisen linssin, valo muuttuu välittömästi tästä hetkestä eteenpäin punaiseksi, ja se välittää tätä merkkiä (punaisuutta), kunnes jokin toinen muutos tapahtuu. Tämä on kausaalinen prosessi, sillä muokkaus yhdessä hetkessä vaikuttaa jatkuvasti muutoksen jälkeenkin, ja prosessi pitää merkin pysyvästi. Kontrastina tälle Salmon (1989, ss. 108–109) esittää ajatuksen pseudoprosesseista, jotka eivät pysty siirtämään merkkejä vastaavalla tavalla. Tällainen tilanne syntyy esimerkiksi, mikäli pitäisimme taivaalla punaista kangasta siten, että majakan pyörivä valo osuu siihen vain hetkellisesti – valo on valkoinen ennen kuin se tulee kankaan kohdalle, muuttuu punaiseksi kankaan kohdalla, ja muuttuu välittömästi takaisin valkoiseksi ohitettuaan kankaan. Tällainen prosessi ei pysty siirtämään merkkiä (punaisuutta) pysyvästi, joten sen ei nähdä olevan kausaalinen.

Salmonin mukaan prosessit, jotka pystyvät siirtämään merkkejä, pystyvät siirtämään informaatiota ja täten myös kausaalista vaikutusta. Kausaalisten prosessien kohdatessa molemmissa prosesseissa tapahtuu muutos, joka on pysyvä, ja jota ei olisi tapahtunut ilman prosessien kohtaamista. Tyypillinen esimerkki on kahden auton törmäys: autojen liikkeet muodostavat kaksi jatkuvaa fyysistä prosessia, ja törmäys muokkaa näitä

kumpaakin tekemällä lommot molempiin autoihin. Kummallakin autolla (prosessilla) on kyky siirtää tätä lommoa (merkkiä) pysyvästi eteenpäin: kun autot lopulta törmäyksen jälkeen lähtevät liikkeelle, lommo kulkee mukana. (Woodward, 2003, s. 351.) Prosessi voi myös olla kausaalinen, vaikka molemmat autot olisivat vahingoittumattomia: olennaista on se, että tapahtuma olisi voinut siirtää merkin. Tällainen tapahtuma ei kuitenkaan olisi olennainen selityksen kannalta. Ajatus olennaisuudesta muodostuu kuitenkin ongelmaksi CM-mallille. Ehkäisy pillerien syönte on kausaalinen prosessi sekä herra että rouva Jonesin kohdalla – syöminen aiheuttaa fyysisiä muutoksia sekä pillerissä että sen syövässä henkilössä. Raskauden todennäköisyyden kannalta tämä on kuitenkin relevanttia ainoastaan rouva Jonesin kohdalla. Pelkkä kausaalisen prosessin havaitseminen ei ole riittävää selityksen kannalta, ja CM-malli ei yksin pysty ratkaisemaan, mitkä merkit (ja minkä merkkien siirtyvyys) on selityksen kannalta olennaista. (Woodward, 2017.)

Tulevan taloustieteitä koskevan luvun kannalta CM-mallin suurimmat ongelmat koskevat kuitenkin sen vaatimusta kausaalisten prosessien seuraamisesta, ja sitä, kuinka tämä voi edes teoriassa onnistua monimutkaisten systeemien kohdalla. Suuri osa taloustieteen selityksistä ei sisällä minkäänlaista fyysistä jatkuvuutta, vaikka sellainen voidaan periaatteessa nähdä tapahtumien taustalla. (Woodward, 2003, ss. 354–355.) Monopoliesimerkin tapauksessa vähentynyt myynti voidaan ymmärtää valtavana määränä yksittäisiä fyysisiä prosesseja, joissa raha tai energia siirtyy paikasta toiseen, mutta tällaisen selityksen antaminen ei tunnu olevan lainkaan hyödyllistä. Se on suorastaan päinvastaista taloustieteen yleisten periaatteiden kannalta: yksinkertaistettujen mallien arvo perustuu juuri niiden kykyyn tavoittaa selitettävän tapahtuman kannalta keskeiset seikat. Monopolin seurausten kannalta on täysin irrelevanttia seurata kaikkia niitä prosesseja, jotka tämän seurauksen reaalisessa maailmassa aiheuttavat. CM-malli kärsii siis merkkien valintaan liittyvien ongelmien lisäksi analyysin tason valintaan liittyvistä ongelmista: monimutkaisissa tapauksissa se tuntuu jopa vievän meitä aivan väärään suuntaan.

Seuraavan luvun tavoitteena on osoittaa, että manipulaatiomalli soveltuu erinomaisesti käsittelemään malleja, joiden tavoitteena on pelkistää todellisuutta, kunnes selitettävän

tapahtuman kannalta keskeiset vaikuttajat on selvillä. Tavoitteenani on lisäksi osoittaa, että manipulaatiomalli voidaan löytää implisiittisesti monen taloustieteen selityksen pohjalta, ja että taloustieteen yleinen näkemys kausaatiosta ja selityksistä on erinomaisesti linjassa manipulaatiomallin kanssa.

4 Selitys taloustieteessä

Kun puhutaan tieteellisestä selittämisestä filosofisessa mielessä, taloustieteen selitykset voidaan nähdä osana yleisempää sosiaalitieteiden selityksiä koskevaa keskustelua. Historiallisesti sosiaalitieteiden selitykset ovat olleet ongelmallisia, koska niiden tapauksessa luonnonlaeista tai odotettavuudesta puhuminen tuntuu menettävän mielekkyytensä. Hempel kohtasi tämän ongelman tutkiessaan historiallisia selityksiä, jotka eivät juuri koskaan viittaa lakeihin. Hempelin mukaan epätäydellisten historiallisten selitysten takana piilee aito DN-mallinen selitys, mutta luonnontieteistä poiketen on helppo huomata, että sosiaalitieteiden tapauksessa vaadittavan lain muotoileminen johtaa jo teoriassa omituisiin lopputuloksiin. Von Wright (1973, ss. 24–25) antaa tästä esimerkin: kuvitellaan tilanne, jossa kuninkaan vaipuminen epäsuosioon selitetään sillä, että hän harjoitti kansan tahdon vastaista politiikkaa. Emme selvästikään voi selittää tätä hyvin kattavan lain kautta, sillä se olisi varmasti vain näennäinen laki: kaikki kansan tahdon vastaista politiikkaa harjoittavat hallitsijat eivät ole epäsuosittuja. Saadaksemme rakennettua aidosti tilanteen selittävän lain, sen on sisällettävä kaikki kyseiseen tilanteeseen vaikuttaneet tekijät. Täten päädymme lakiin, jonka mukaan kaikki kuninkaas, jotka harjoittavat täysin samoissa olosuhteissa täysin samanlaista politiikkaa kuin tämä kuningas, vaipuvat epäsuosioon. Tässä muodossa laki voi kuitenkin selittää vain yhden tapahtuman, ja se on nurinkurisesti sama tapahtuma, jonka pohjalta laki on rakennettu. Laki ei siis ole yleinen sanan missään merkityksessä.

Taloustiede on täynnä samanlaisia tilanteita. Ongelma muodostuu siitä, että emme voi tarkastella perimmäiseltä olemukseltaan sosiaalisia tapahtumia erillään niitä ympäröivistä tekijöistä. Tästä syystä parhaatkin lainomaiset väitteemme, kuten ”koulutus kasvattaa tulotasoa” ja ”tuotteen hinnan nostaminen laskee sen kysyntää”, eivät täytä tiukan lainomaisuuden määritelmää. Tästä huolimatta kaikissa kattavissa otoksissa voidaan

havaita selvä yhteys esimerkiksi väestönosien koulutus- ja tulotason välillä, ja haluaisimme varmasti väittää näiden kahden välillä olevan kausaalinen yhteys (vaikka tämä yhteys perustuisi muihin koulutuksen mukanaan tuomiin mekanismeihin). Kuten olemme jo huomanneet, invarianssin käsite soveltuu erinomaisesti tällaisten tilanteiden käsittelyyn: esimerkiksi väite ”koulutus kasvattaa tulotasoa” on siinä määrin invariantti, että meidän ei usein edes tarvitse täsmentää tarkemmin, mitä kontroleja on oltava paikallaan. Väite kuvaa riittävän todenmukaisesti niin montaa erilaista aikaa ja paikkaa, että koulutus selittää tulotasoa lähes näistä riippumatta – tästä syystä se myös vaikuttaa lainomaiselta, vaikka tosiasiaa kyse on vain yleisyydestä siinä mielessä, missä se invarianssin yhteydessä ymmärretään.

Pitkästi DN-mallin vaatimien lakiväitteiden vuoksi sosiaalitieteiden saralla on myös väitely siitä, onko selitys välttämätöntä ymmärryksen kasvattamiseksi. Kilpailevan näkemyksen mukaan *tulkintojen* tulisi olla sosiaalitieteiden keskiössä: inhimillinen toiminta ei mukaudu universaaleihin lakeihin, vaan se perustuu uskomuksiin ja haluihin, joita voidaan selitysten sijaan tutkia esimerkiksi empatian avulla. Täten inhimillisen toiminnan ymmärtämisen kannalta olennaista ei ole toiminnan selittäminen, vaan sen tulkitseminen. (Kuorikoski, 2010b, ss. 21–22). Taloustiede tuntuu tässä mielessä olevan erilainen muihin sosiaalitieteisiin nähden, sillä selitys, ei tulkinta, on aina ollut taloustieteen keskiössä. Friedmanin (1953, ss. 7–9) klassisen näkemyksen mukaan taloustieteen päällimmäinen tavoite on tuottaa ennustuksia, ja näiden ennusteiden voidaan suoraviivaisesti nähdä rakentuvan selitysten varaan – aivan kuin Hempelillä, selitys kertoo, mitä voimme odottaa tapahtuvan. Samalla kun selitysmallimme ovat liikkuneet kohti kausaatiota, myös taloustiede on siirtynyt ennusteista kausaalisten mekanismien etsimiseen – moderneissa teksteissä on yleistä huomata sanoja ”aiheuttaa” ja ”selittää” käytettävän synonyymisesti. Samoin näkemykset siitä, että taloustieteellinen mallintaminen on selittävää toimintaa, ja että selitys koostuu kausaalisten mekanismien ilmaisemisesta, ovat nykyään arkipäivää.

Tässä luvussa pyritään Uskali Mäen teoriaa hyödyntämällä antamaan kuvaus siitä, miten taloustiede tutkii monimutkaisia sosiaalisia tapahtumia koskevia kausaalisia

mekanismeja. Kuten jo johdannossa lyhyesti todettiin, *mallintaminen* on ehdottomasti taloustieteen tärkein työkalu tämän tavoitteen saavuttamiseen.

4.1 Mallit sijaissysteeminä

Mallit ovat tiettyä tarkoitusta varten rakennettuja kuvauksia jostain kohteesta, esimerkiksi teoriasta tai datajoukosta. Tämä mielessä pitäen mallien rakentamisessa ei ole kyse ainoastaan siitä, pystymmekö jollain mallilla M kuvaamaan jotain kohdetta K , vaan lisäksi, että pystymmekö tämän kuvauksen avulla saavuttamaan selityksellisen päämäärämme. Tämä ajatus on intuitiivisesti tuttu kaikille, ja tilannetta voidaan verrata arkielämässä usein tehtäviin analogioihin: osa analogioista on osuvia ja tavoittavat hyvin kohteensa keskeiset puolet, kun taas toiset syystä tai toisesta epäonnistuvat täysin – melkein mikä tahansa voi toimia kuvauksena mille tahansa muulle, mutta onnistuminen on riippuvaista ulkoisista tekijöistä. Kuten analogioiden, myös mallien tulee olla käyttötarkoitukseensa sopivia monilla eri mittareilla. Malli on *malli* juuri käyttötarkoituksensa vuoksi. Mäki (2009b, s. 75) määrittelee tämän seuraavasti: ”agentti A käyttää mallia M kuvaksena kohdesysteemistä K päämäärän P vuoksi yleisölle Y siten, että se herättää aitoja kysymyksiä samankaltaisuudesta $M:n$ ja $K:n$ välillä, ja kommentaarin C avulla huomioi näitä tekijöitä ja koordinoi niiden suhteita.” Mallinnuksessa on onnistuttu, mikäli valittu työkalu on tehtävänsä sopiva. Korostetun tärkeää on se, että tämä muotoilu ei vaadi mallilta täydellistä kuvausta kohteestaan – riittää, että mallia hyödyntämällä voidaan nostaa esiin kuvauksen ja kuvattavan välisiä yhteyksiä. Onnistunut kuvaus kuvaa kohdettaan halutulla tavalla, jonka määrittää tavoiteltu päämäärä ja vastaanottava yleisö.

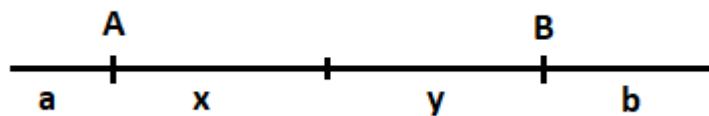
Mäki (2009b, s. 76) näkee mallit sijaissysteeminä, joita käsittelemällä voimme saada suoraa tietoa mallin sisäisestä maailmasta. Mäen mukaan kuvauksellisuutensa vuoksi mallien avulla voidaan kuitenkin saada myös epäsuoraa tietoa niiden kuvaamasta aidosta maailmasta – hyvä malli toimii vain sijaisena todelliselle systeemille. Tässä näkemyksessä mallien ”epätotuus” johtuu siitä, että mallit eivät vaikuta täyttävän onnistuneen kuvauksen kriteerejä: ne eivät ole tarpeeksi samankaltaisia todellisen maailman kanssa, jotta hyppy mallin sisäisestä tiedosta mallin ulkopuoliseen maailmaan

olisi oikeutettu. Kun yhteys todelliseen maailmaan katkeaa, mallit muuttuvat sijaisysteemeistä täysin korvaaviksi systeemeiksi. Korvaavalla systeemillä Mäki tarkoittaa mallia, jota ei enää voida nähdä vain sijaisena tarkasteltavalle maailmalle, vaan joka on itse muuttunut tarkastelun kohteeksi, ja jota tarkastelemalla voidaan saada tietoa ainoastaan mallista itsestään. Mitä enemmän malleja yksinkertaistetaan, sitä todennäköisemmäksi tulee mahdollisuus mallin muuttumisesta täysin korvaaviksi systeemiksi, josta vedettyjä johtopäätöksiä ei voida enää hyödyntää sellaisenaan missään aidossa systeemissä.

Esitellään seuraavaksi esimerkki yksinkertaisesta mallista tulevan käsittelyn tueksi, jotta saadaan käsitys siitä, mitä yksinkertaistamisella konkreettisesti tarkoitetaan.

4.1.1 Hotellingin malli

Klassisessa artikkelissaan Hotelling (1929) esittää mallin, joka pyrkii selittämään nykyään Hotellingin lakina tunnettua periaatetta, jonka mukaan monilla markkinoilla tuottajien on järkevää tehdä tuotteistaan täysin samankaltaisia, vaikka tämä ei ole sosiaalisesti optimaalinen lopputulos. Yksinkertaisesti esitettynä malli on seuraava: kuvitellaan kaksi kilpailevaa liikettä, A ja B, täysin suoran kadun varrella, ilman minkäänlaista maantieteellistä vaihtelua. Täysin identtiset asiakkaat ovat tasaisesti jakautuneet tämän kadun varrelle, ja myös liikkeet ovat täysin identtisiä kaikin muin tavoin, paitsi että ne sijaitsevat kadun vastakkaisissa päissä. Asiakkailla liikkeessä käyntiin liittyy matkustuskustannuksia, jotka riippuvat näiden etäisyydestä liikkeeseen, joten jokainen asiakas maksimoi hyötyään suhteessa kuljettavaan matkaan ja tuottajan pyytämään hintaan.



Hotellingin lineaarisen kaupungin malli

Lähtökohtaisesti mallin mukaan liike A saa kaikki asiakkaat osuudelta $a + x$ ja liike B osuudelta $b + y$. Lisäksi $x = y$, ja liikkeiden välisessä keskipisteessä sijaitseva kuluttaja ei välitä kummassa liikkeessä hän käy. Mikäli liike A laskisi hintojaan, x :n pituus kasvaisi ja y :n pienenesi, kunnes saavutettaisiin jälleen piste, jossa alentunut hinta ei korvaa ylimääräisestä matkasta koituvia kustannuksia. Mallin tärkein opetus liittyy kuitenkin siihen, mitä tapahtuu, jos annamme liikkeille mahdollisuuden siirtyä janalla toiseen pisteeseen. Mikäli liike B toimii ensin, sen kannattaa siirtää itsensä välittömästi A:n oikealle puolelle, jolloin se saa haltuunsa kuluttajat alueelta $x + y + b$. Liikkeet eivät saa siirtyä samaan pisteeseen, joten A:n vastaus tähän on puolestaan nyt siirtyä B:n oikealle puolelle. Tämä jatkuu, kunnes molemmat liikkeet ovat janan keskellä ja jakavat markkinat tasan – saavuttaakseen suurimman mahdollisen markkinaosuuden liikkeet siirtyvät niin lähelle toisiaan kuin pystyvät, muuttumatta kuitenkaan identtisiksi. Sosiaalisesti optimaaliset sijainnit liikkeille löytyisivät neljäsosamatkan päästä kummastakin päädyistä: tällöinkin liikkeet jakaisivat markkinat tasan, ja samalla minimoitaisiin asiakkaiden kulkema matka. Mallin kuvaamat mekanismit kuitenkin estävät tämä saavuttamisen.

Hotellingin mukaan tällaista samankaltaistumisen ilmiötä voidaan havaita niin taloudellisissa kuin ei-taloudellisissa ympäristöissä. Rationaalinen pyrkimys tavoittaa suurempi osa markkinoista selittää, miksi poliittiset puolueet ovat kaikki lopulta hyvin lähellä toisiaan, ja miksi kaikki siiderit maistuvat ja kaikki vaatteet näyttävät samalta. Hotellingin malli tuntuu todella antavan hyväksyttävän selityksen valtavalle määrälle arkisia ilmiöitä, joiden olemassaolon tajuaa kenties vasta mallia opiskeltuaan. Yhtäkkiä on selvää, miksi päivittäistavara-kaupat ilmestyvät aina pareittain, ja miten yhteen kortteliin mahtuu neljä parturikampaamoja. Sen ilmeisestä selitysvuorosta huolimatta malli ei kuvaa ”todennukaisesti” yhtäkään selittämäänsä tapahtumaa.

4.1.2 Kausaalisten tekijöiden eristäminen

Kuten Hotellingin lineaarinen kaupunki, mallit ovat jo määritelmällisesti kuvitteellisia, koska ne syntyvät havaintojen sijaan tutkijoiden mielikuvituksesta. Näitä kuviteltuja maailmoja luodessaan tutkijat sekä talous- että luonnontieteen alalla käyttävät hyvin epärealistisia oletuksia esimerkiksi täydellisestä informaatiosta, kitkattomista tasoista, tai

täysin rationaalisista identtisistä kuluttajista. Kaikille oletuksia käyttäville on selvää, että ne eivät ole todenmukaisia – oletukset eivät ole virheitä, ja niiden mahdottomuus todellisessa maailmassa on sisäänrakennettu ominaisuus. Mäki (2009b, s. 78) puhuu niistä strategisina epätotuuksina, joita käytetään tietyn päämäärän saavuttamiseen: oletusten avulla on tarkoitus eristää jokin tärkeä riippuvuussuhde muilta siihen vaikuttavilta tekijöiltä. Tällaista teoreettisen tason todellisuuden vääristämistä kutsutaan idealisaatioksi (Mäki, 1992, s. 324). Yleisesti taloustieteessä idealisaatio toteutetaan antamalla joillekin muuttujille käytännössä mahdottomia arvoja – oletamme usein esimerkiksi työllisyyden olevan sadassa prosentissa, tai vaihtokustannusten nollassa, vaikka nämä tilanteet ovat käytännössä mahdottomia. Näillä toimenpiteillä pyritään eristämään tietyt kausaaliset mekanismit muilta tapahtumaan vaikuttavilta tekijöiltä.

Kun tarkastelemme pelkästään tämän tyyppistä kausaalisten mekanismien eristämistä, voimme huomata sen ”vääristävän” todellisuutta vain yhdellä tavalla. Kun eristämme onnistuneesti jonkin kausaalisten vaikuttajien joukon kaikilta muilta tekijöiltä, emme varsinaisesti mallinnuksessamme muokkaa maailmaa ja tee siitä täysin epätodenmukaista, vaan jätämme vain tiettyjä asioita huomioimatta. Hotellingin mallissa oletamme pois maantieteelliset tekijät ja sijainnin ulkopuoliset vaikuttimet ostopäätökseen, ja huomio keskitetään siihen, että yhdenmukaistaminen on rationaalinen päätös voittoa maksimoivalle yritykselle. Täten sanoessamme mallin olevan epärealistinen, tarkoitamme, että se kattaa vain suhteellisen pienen osuuden tapahtumaan vaikuttaneista tekijöistä; vastaavasti realismilla tarkoitetaan tässä yhteydessä kattavuutta. Mallin todenmukaisuutta voidaan siis karkeasti arvioida vertaamalla eristetyn joukon ”koon” suhdetta pois suljettujen tekijöiden joukon ”kokoon”. Tämän myötä realismisuus muodostaa jatkuvan skaalan, ja mitä suuremman joukon tapahtumia jokin malli kattaa, sitä realistisempi se on. (Mäki, 1992, s. 321.)

Eristämiseen liittyy usein kuitenkin myös toinen aspekti, jolla voidaan nähdä olevan suurempi vääristävä vaikutus mallin ja todellisuuden suhteeseen. *Abstraktiossa* yleiset havainnot pyritään eristämään yksittäisiltä tapahtumilta. Tämä on hyödyllistä, jos haluamme puhua jostain yleisemmästä ilmiöstä ottamatta kantaa siihen liittyviin erilaisiin yksityiskohtiin: näin tapahtuu esimerkiksi puhuttaessa ”markkinoista” yleisellä tasolla.

(Mäki, 1992, s. 322.) Makrotaloustieteessä kulutushyödykkeet aggregoidaan usein yhdeksi yleiseksi kulutushyödykkeeksi, jolloin voimme tarkastella kaikkien kulutushyödykkeiden markkinoita abstraktimmalla tasolla välittämättä yksittäisten markkinoiden ominaisuuksista. Siirtymällä vielä abstraktimmalle tasolle pääsemme käsiksi yleiseen ajatukseen markkinasta, jota voidaan soveltaa myös muihin kuin kulutushyödykkeisiin. Mitä abstraktimpi malli, sitä epärealistisempi se on, ja sitä heikommin se kuvaa mitään tiettyä konkreettista tapahtumaa; vastaavasti konkreettisemmat mallit, jotka sisältävät enemmän yksityiskohtia, ovat realistisempia.

Näillä tekijöillä on mielenkiintoisia yhteyksiä selityksellisen voiman ulottuvuuksiin. Kun tarkennamme explanandumia, se usein tekee selityksestämme herkemmän. Vastaavasti integraation aste ja kognitiivinen helppous usein kärsivät siitä, mitä faktuaalisesti tarkempi selitys on. (Ylikoski & Kuorikoski, 2010, s. 212.) Näin ollen idealisaatiossa on usein kyse siitä, että haluamme selityksemme olevan yksinkertaisempi ja paremmin integroitunut johonkin yleiseen teoriaan, ja abstraktiossa pyrimme tekemään selityksestä epäherkemmän explanandumin tarkkuuden kustannuksella. Yksinkertaisiin selityksiin liitettävät negatiiviset miellelyhtymät ovat vain toinen puoli siitä kolikosta, jonka selityksen vahvuus muodostaa: yksinkertaiset mallit parantavat merkittävästi selityksellistä voimaa tietyillä ulottuvuuksilla, ja heikentävät sitä toisilla. Tutkijan tehtävä on päättää, kuinka tämä tasapainottelu parhaiten maksimoi selityksellistä voimaa.

4.1.3 Teoreettiset eristykset

Edellä kuvattu pyrkimys relevanttien kausaalisten mekanismien eristämiseen nostaa esiin houkuttelevan vertauksen mallien ja empiiristen kokeiden välillä. Hyvät empiiriset kokeet sulkevat pois tiettyjä aitoja tapahtumaa vaikuttavia tekijöitä eristäen tarkasteltavat muuttujat muusta maailmasta. Laboratoriokokeissa pyritään usein juuri tähän. Nämä aitoihin tapahtumiin liittyvät ja konkreettisiin toimin syntyneet *materiaaliset eristykset* luovat kontrastin taloustieteen *teoreettisille eristyksille*. Teoreettiset eristykset eivät toteudu aidossa maailmassa, vaan tutkijan mielessä, jossa hän mallien ja ajatuskokeiden avulla pyrkii erottelemaan aidosta maailmasta relevantit seikat olettamalla tietyt muuttujat idealisaation avulla merkityksettömiksi. Kokeellisessa taloustieteessä pystytään jonkin

verran hyödyntämään materiaalista eristämistä, mutta lähtökohtaisesti taloustieteessä laajojen kontrolloitujen kokeiden järjestäminen on äärimmäisen vaikeaa ja kallista. Valtiontalouden tasolla kokeet voidaan nähdä käytännössä mahdottomina. Monien tutkimuskysymysten kohdalla taloustieteen ainoa vaihtoehto on täten tukeutua teoreettiseen kokeiluun mallien avulla. (Mäki, 1992, ss. 325–326.)

Teoreettisen eristämisen saavuttamiseksi taloustieteilijöillä on käytössään kaksi työkalua: idealisaatio ja poisjätto (*omission*). Kuten jo aiemmin todettiin, idealisaatiossa jokin muuttuja rajoitetaan johonkin tiettyyn (tarpeen vaatiessa jopa mahdottomaan) arvoon, jotta sen vaikutus lopputulokseen voidaan poistaa. Näin ollen idealisaatio on edelleen kuvaus kohteestaan: se kuvaa aitoa tekijää, mutta antaa sille arvon, joka ei välttämättä voi aidossa maailmassa koskaan tulla todeksi. Idealisaatiot ovat siis määritelmällisesti epätosia kuvauksia. Kun esimerkiksi jonkin tasapainomallin kuvaama systeemi kohtaa shokin, oletamme, että vaikutus on välitön – toisin sanoen annamme mukautumisnopeutta kuvaavalle muuttujalle äärettömän suuren arvon. Näin voimme jättää analyysistä pois taloudelliseen toimintaan liittyvän jäykkyyden, ja tarkastella shokin vaikutuksia ilman tutkittavan ilmiön kannalta turhaa viivettä. Mukautumisaikaa kuvaava muuttuja on kuitenkin huomioitu ja kuvattu mallissa: se on poimittu tärkeänä, ja sille on annettu tarkoituksellisesti epätosi arvo, jotta malli tavoittaa paremmin päämääränsä. (Mäki, 1992, ss. 328–329.)

Poisjätto poikkeaa idealisaatiossa siinä, että poisjätettävää muuttujaa ei kuvata mallissa lainkaan. Ajatus on, että poisjätettävän muuttujan vaikutus lopputulokseen on niin pieni, ettei sitä ole syytä kuvata tai sen arvoa rajoittaa. Poisjätettävien muuttujien joukkoon voi kuulua esimerkiksi lakitekniisiä tai logistisia seikkoja, mutta niihin kuuluu myös loputon määrä täysin irrelevantteja asioita, kuten esimerkiksi ruohonkorsien määrä Suomen Pankin pihalla. (Mäki, 1992, s. 330.) Pelkässä matemaattisessa mallissa ei ilmaista poisjätettyjä muuttujia, joten poisjätto on äärimmäisen tehokas teoreettisen eristämisen keino: mallin sisäisessä systeemissä poisjätetyt muuttujat eivät voi vaikuttaa tarkasteltaviin muuttujiin lainkaan. Poisjätetyt muuttujat nousevat yleensä esiin vasta mallin ympärille muodostuvassa kommentaarissa, joka kertoo, miten mallia tulee tulkita, ja mitä sillä pyritään osoittamaan. Pelkkä muutamista funktioista koostuva matemaattinen

malli on itsenäään hyvin ohut kuvaus mistään tapahtumasta, joten kommentaarissa luodaan matemaattiselle formalismille konteksti, joka voi vaihtua yleisön mukaan. Tässä kommentaarissa huomioidaan yleensä asioita, joita malliin ei ole sisällytetty, ja ne idealisoidaan vasta ikään kuin mallin ulkopuolella. (Mäki, 1992, ss. 330–331.) Suurin osa poisjätetyistä muuttujista jätetään kuitenkin pois hiljaisesti niiden täyden irrelevanssin takia.

4.1.4 Mallit ovat kokeita ja kokeet ovat malleja

Mallinnuksessa ja sen onnistumisen arvioinnissa on otettava huomioon niin semanttisia, epistemologisia kuin pragmaattisia seikkoja (Mäki, 2005, s. 305). Miten hyvin malli kuvaa tavoittelemaansa kohdetta? Onko kuvaus riittävän samankaltainen kohteensa kanssa, jotta sen avulla voidaan nostaa esiin relevantteja samankaltaisuuden liittyviä kysymyksiä? Pystymmekö mallin avulla saamaan epäsuorasti tietoa myös sen kuvaamasta kohteesta? Entä onko malli sopiva tavoitteeseensa ja kohdeyleisöönsä nähden? Nämä kysymykset ja niiden vastaukset määrittävät mallinnuksen onnistumista, joten itse mallien ontologia jää hyvin joustavaksi. Lähes mikä tahansa voi toimia mallina mille tahansa muulle, niin kauan kun se hyväksyttävästi täyttää mallinnukselta vaaditut ehdot (Mäki, 2005, s. 305). Keskeinen mallinnuksen ominaisuus on todellisuuden yksinkertaistaminen, joka saadaan aikaiseksi eristämällä tiettyjä relevantteja seikkoja monimutkaisista tapahtumista.

Kuten ehkä on ilmeistä, kaikki edellä mainitut kysymykset ovat aiheellisia myös empiiristen kokeiden kohdalla. Kun laboratoriossa luodaan koe, joka pyrkii testaamaan jotain laboratorion ulkopuolella tapahtuvaa ilmiötä, tämä koe voidaan käytännössä nähdä mallina todellisen maailman tapahtumasta. Täten sen onnistumisen arvioinnissa on käytettävä samanlaisia kriteerejä kuin täysin teoreettisten mallien kohdalla. Piinaavan ajankohtainen esimerkki koskee virusten säilymistä erilaisilla pinnoilla. Laboratoriokokeissa on havaittu koronaviruksen säilyvän tietyillä pinnoilla jopa kolmen päivän ajan, mutta kaikki terveystieteelliset huomioivat, että tästä ei voida vielä tehdä lopullisia johtopäätöksiä viruksen säilymisestä arkiympäristössä. Mistä tämä epävarmuus voi johtua? Mäen näkemyksessä voimme pitää laboratoriokoetta mallina todellisesta

tilanteesta, jossa virus on tarttunut johonkin pinnalle – tässä mallissa on eristetty pois kaikki arkielämän aiheuttamat muuttujat esimerkiksi pintojen laadusta, lämpötilasta, kosteudesta ja viruksen pinnalle päätyminen tavasta. Voimmekin nyt aiheellisesti kysyä, onko koe tarpeeksi samankaltainen kuvaamansa tilanteen kanssa, ja voimmeko täten saada sen avulla epäsuorasti tietoa myös arkielämän tilanteesta. Selvästi voimme saada jotain tietoa, joka on sovellettavissa ainakin siihen pieneen osaan todellisuudesta, jonka kokeen muodostama malli maailmasta eristää. Täysi yleistäminen on kuitenkin kontrollien paikallaan pysyessä mahdotonta – jotta voisimme selittää asiaa laajemmin, meidän täytyy muuttaa koetta vastaamaan erilaisiin kysymyksiin. Kokeet ovat siis samalla tavoin pragmaattisia kuin teoreettiset mallit.

Teoreettiset mallit ja empiiriset kokeet ovat monilta tavoin verrattavissa toisiinsa: molemmat luovat kuvauksen, joka eristää tiettyjä suhteita, joita voidaan sitten tarkastella manipulaatioiden avulla. Niiden keskeinen ero on, että kokeissa manipulaatiot toteutetaan konkreettisesti, kun taas malleissa ne saadaan aikaiseksi teoreettisesti. Toisaalta tämä myös mahdollistaa paljon tiukempien eristysten tekemisen teoreettisissa malleissa, kun oletuksia voidaan tehdä välittämättä fyysisistä rajoitteista. (Mäki, 2005, ss. 311–312.) Tämän näkemyksen mukaan väite, jonka mukaan ”taloustieteelliset mallit ovat epätosia” on yhtä mieletön kuin ”empiiriset kokeet ovat epätosia.” Mallit luovat yksinkertaistetun sijaisjärjestelmän tiettyä tarkoitusta varten, ja tämän tarkoituksen kannalta niiden ei edes tule olla ”tosia”, mikäli tällä tarkoitetaan täydellistä kattavuutta. Oman näkemykseni mukaan mallien avulla ei siis pyritä tuottamaan välttämättä tarkkaa ja todenmukaista kuvausta aidosta maailmasta, vaan niiden avulla pyritään tavoittamaan kausaalisia rakenteita Woodwardin tarkoittamassa mielessä.

4.1.5 Sijaisjärjestelmien kontrafaktuaalinen uskottavuus

Vaikka mallit eivät ole tosia tai epätosia sanojen tiukassa merkityksessä, on silti aiheellista kysyä, miten mallien avulla voidaan tehdä päätelmiä, jotka pätevät myös niiden ulkopuolisessa maailmassa. Tähän pohjautuen Sugden (2000) esittää mahdollisen haasteen Mäen teorialle. Sugdenin mukaan eristämiseen perustuvien mallien tuloksia voidaan yleistää aitoon maailmaan vain, jos eristämisen keinot ovat hyväksyttävissä

empiirisiä yleistyksiä näiden muuttujien toiminnasta. Mikäli eristykset pohjautuvat hyväksyttäviin yleistyksiin, voimme deduktiivisesti päätellä, missä määrin mallin tulokset vaikuttavat myös näiden yleistysten taustalla olevien yksittäisten ja aitojen tapahtumien kohdalla. Toisaalta, mikäli eristyksemme *eivät* ole hyväksyttäviä empiirisiä yleistyksiä, mallillamme ei ole empiiristä pohjaa, eikä deduktiivinen päättely mallin sisäisestä aitoon tilanteeseen ole mahdollista.

Ovatko eristykset useinkaan hyväksyttäviä empiirisiä yleistyksiä? Sugdenin (2000, s. 17) mukaan eivät – täysin rationaalinen hyötyä maksimoiva kuluttaja ei ole hyväksyttävä yleistys yhdestäkään ihmisestä. Tästä seuraa, että mallien sovellusala on välttämättä niiden tekemien oletusten rajoittama: mallit eivät ole pelkkiä yksinkertaistuksia maailmasta, vaan niihin on myös lisätty jotain, mitä aidosta maailmasta puuttuu. Deduktiivisen päättelyn sijaan Sugden ehdottaa näin ollen mallinnuksen pohjaksi induktiivista päättelyä: mallien tehtävä ei olekaan yksinkertaistaa maailmaa sijaissysteemien avulla, vaan *luoda* vaihtoehtoinen maailma, jonka pohjalta voimme tehdä aitoa maailmaa koskevia induktiivisia päätelmiä. (Sugden, 2000, s. 19.) Induktiivinen päätelmä on oikeutettu, jos luomamme maailma on kontrafaktuaalisesti uskottava. Mitä helpommin voimme nähdä mallin kuvauksena siitä, millainen maailma *voisi* olla, sitä parempi oikeutus meillä on uskoa, että kun näemme mallissa F:n aiheuttavan G:n, niin sama mekanismi on toiminnassa myös aidossa maailmassa, kun havaitsemme nämä tekijät yhdessä (Sugden, 2000, s. 24).

Luomiseen perustuvassa näkemyksessä Hotellingin malli ei perustu asioiden poistamiseen todellisesta maailmasta, vaan kyse on kuvitteellisen maailman luomisesta, joka on relevanteilta osa-alueilta verrattavissa aitoon maailmaan. Tässä maailmassa kaupungeissa ei ole maantieteellistä vaihtelua, ja asiakkaat valitsevat liikkeet vain niiden sijainnin perusteella. Emme tarvitse paljoa mielikuvitusta pystyäksemme kuvittelemaan tällaisen maailman mahdollisuutta ja olemassaoloa: se on uskottava, koska se voisi teoriassa olla aito. Cartwright (2009, s. 46) antaa erittäin osuvan minimalistisen tulkinnan maailman uskottavuudesta: uskottava maailma sisältää tekijöitä, jotka kuuluvat aitoon maailmaan, ja jotka käyttäytyvät systemaattisesti aidon maailman rajoitusten mukaisesti.

Hotellingin malli toimii, tulkitsimme sitä sitten eristämiseen tai luomiseen perustuvan teorian kautta. Tämän lisäksi tuntuu, että emme ole tosiasiasa vaikuttaneet lopulliseen tulkintaan lähes ollenkaan: ero Sugdenin ja Mäen teorioiden välillä jää melko näennäiseksi. Sen sijaan, että ne tarjoaisivat kaksi täysin erillistä näkemystä, ne tuntuvat ainoastaan korostavan erilaisia, toisiaan tukevia näkökulmia liittyen mallintamiseen. Ennen kaikkea Mäki (2009a, ss. 31–32) näkee Sugdenin ymmärtävän väärin eristämisen todellisen merkityksen ja tavoitteen. Eristämisen tarkoitus on ainoastaan sulkea jokin systeemi tiettyjen tekijöiden vaikutukselta ja se, miten tämä toteutetaan, ei ole varsinaisesti merkityksellistä – joskus eristämisen aikaansaaminen saattaa vaatia aidosta maailmasta puuttuvien asioiden lisäämistä malliin, mutta tämä ei lopullisen tavoitteen kannalta ole ongelma. Eristäminen ei ole metodi mallien luomiseen, vaan se on mallien lopputuote: rakennamme malleja siitä syystä, että haluamme eristää jonkin tekijän tietyiltä muilta vaikutuksilta. Tämä tarkoittaa, että mallintamista ei voida nähdä pelkästään luomisena tai yksinkertaistamisena, vaan *sekä* luomisena *että* yksinkertaistamisena. Mallintamisessa pyritään vain luomaan sijaisysteemi, jossa vaadittavat eristykset saadaan aikaan.

Mikä ero uskottavan maailman ja sijaisysteemin välillä lopulta on? Kysymys vaikuttaa täysin tulkinnanvaraiselta, sillä uskottavat maailmat voidaan hyvin helposti nähdä sijaisysteeminä ja päinvastoin (Mäki, 2009a, ss. 34–35). Uskottavat maailmat ovat kuvauksia, joita käytetään aidon maailman sijasta. Ne sisältävät tärkeitä tekijöitä aidosta maailmasta ja jättävät muut tekijät huomiotta. Lisäksi ne ovat sitä parempia, mitä osuvammin niiden avulla voidaan tarkastella kysymyksiä samankaltaisuudesta aidon maailman ja mallin välillä. Sugdenin teoria lähinnä tarjoaa meille yhden ulottuvuuden lisää, jolla arvioida sijaisysteemin toimivuutta: kontrafaktuaalinen uskottavuus vaikuttaa olevan olennainen osa sitä, mikä tekee relevanttien seikkojen arvioinnin mallin ja todellisuuden välillä mahdolliseksi.

4.2 Manipulaatioteoria ja taloustieteellinen mallinnus

Mallien avulla pyritään eristämään kausaalisia suhteita ja mahdollistamaan eristettyjen suhteiden tarkastelu manipulaatioiden avulla – tämän näkemyksen yhteys Woodwardiin

on melko ilmeinen. Mäki ei itse ole eksplisiittisen woodwardilainen, mutta hän allekirjoittaa manipulaatioteorian näkemyksen selityksistä teoriansa yhteydessä lähes sanasta sanaan. Tämä kuvaus on niin tarkka, että se on syytä esittää suorana sitaattina:

”Useiden taloustieteen mallien ei ole tarkoitus vastata ”miksi” kysymyksiin: niitä ei käytetä selittämään tiettyjä tapahtumia kertomalla, miksi ne tapahtuivat. Tämän sijaan niitä käytetään vastaamaan kysymyksiin siitä, miksi jokin tapahtuma *olisi voinut* tapahtua. Sen sijaan että kysyisimme ”Miksi P tapahtui?”, kysymmekin ”Miksi P olisi voinut tapahtua?” siten, että explanandum P on jokin yleistys sen sijaan, että se olisi tietty tapahtuma. [...] Malli ei mahdollista päättelyä parhaaseen selitykseen. Se mahdollistaa vain päättelyn *mahdolliseen* selitykseen: se tuo esiin mahdollisen mekanismin, joka olisi voinut tuottaa sen aggregoidun lopputuloksen (tai heikommin: olisi voinut osallistua sen tuottamiseen), joka havaitaan. (Mäki, 2009a, s. 38.)”

Mallien tehtävä on siis ehdottaa selityskandidaatteja. Kuten mustan korpin tapauksessa, selityksen tulisi koostua niistä väitteistä, jotka kertovat, miten korpin väriin voidaan vaikuttaa. Samoin taloustieteelliset selitykset koskevat tapahtumaan vaikuttavia mekanismeja – meitä ei niinkään kiinnosta se kausaalinen historia, joka tapahtumaan on suoraviivaisesti johtanut, vaan se kausaalinen *rakenne*, josta selitettävä tapahtuma voi nousta esiin. Sitaatissa tulee myös epäsuorasti esiin taloustieteen pyrkimys selitykselliseen epäherkkyyteen ja integraatioon, tietenkin explanandumin ja faktuaalisen tarkkuuden kustannuksella. Mielestäni Mäen mallintamisen teorian on tarkoitus valottaa sitä, miten mallien avulla voidaan päästä käsiksi selityksiin siten, kuten sitaatti ilmentää. Lisäksi väitän, että tämä selityksen tyyppi on täysin Woodwardin esityksen mukainen ja voidaan tavoittaa manipulaatioteorian mukaisten kausaalisten suhteiden kuvauksella. Toisin sanoen, manipulaatioteoria antaa merkityksen mallinnuksen aikaansaamille eristyksille ja tämän prosessin tuottamille havainnoille.

Mäen teorian keskeiset termit ja ajatukset ovat taloustieteen tarpeisiin sopivia käytännön sovelluksia Woodwardin ajatuksista. Nämä yhteydet ovat laajoja ja niitä on useita.

4.2.1 Mahdolliset mekanismit ja syiden monimuotoisuus

Mallintamisessa pyritään luomaan systeemi, joka koostuu selitettävästä tapahtumasta ja jostain joukosta selittäviä muuttujia. Lopullinen tavoite on se, että mallin kuvaamat mekanismit itsessään tuottaisivat selitettävän tapahtuman: sen tulee nousta esiin luomastamme kausaalisesta rakenteesta. Hyvä malli kuvaa todenmukaisesti tätä rakennetta, ja sen avulla voidaan nähdä, miten lopputulos riippuu johdonmukaisesti esitetyistä syistä. Hotellingin mallin kohdalla tämä rakenne muodostuu esimerkiksi matkustuskustannuksista, kuluttajien rationaalisuudesta ja yritysten voittojen maksimoinnista. Kun olemme sisällyttäneet nämä elementit malliin, voimme niitä manipuloimalla tutkia, millaisia vaikutuksia niillä on lopputulokseen ja tarkastella, miten samankaltaistuminen seuraa näiden muuttujien ja mekanismien toiminnasta. Näin ollen mallit voivat vain paljastaa mahdollisia mekanismeja: niiden avulla voidaan tutkia, miten jokin tapahtuma *voisi* syntyä. Mallit eivät kuitenkaan itsessään kerro mitään siitä, ovatko nämä mekanismit todella toiminnassa tietyissä aidoissa tilanteissa. Tämän arviointi on mallin ulkopuolinen kysymys.

Jotta mahdollinen mekanismi todella vaikuttaa selitettävään tapahtumaan, sen on oltava tapahtuman kausaalinen syy jollain manipulaatioteorian määräämällä tavalla. Aidon mekanismin on siis täytettävä joko (**M**):n tai (**KS**):n määrittämät ehdot. Malliin valitut muuttujat määrittävät joukon V , jonka suhteen mahdollisen mekanismin kausaalista vaikutusta arvioidaan. Kuten olemme huomanneet, muuttujajoukon V valinta vaikuttaa olennaisesti siihen, näemmekö jonkin tapahtuman toisen syynä vai emme. Tästä on osittain kyse myös arvioidessamme sitä, ovatko mallit sovellettavissa myös aitoon maailmaan: mikäli mallista puuttuu jokin relevantti muuttuja, voimme päätyä arvioimaan tarkasteltavan mekanismin kausaalista vaikutusta täysin väärin.

4.2.2 Interventiot ja eristykset

Arvioidessamme täyttääkö jokin mekanismi kausaalisen syyn ehdot, meidän on suoritettava interventioita valittuun muuttujajoukkoon. Ottaaksemme esimerkki suoran syyn tapauksesta, tämä vaatii kaikkien muiden V :n muuttujien kontrolloinnin tiettyihin arvoihin. Mäen terminologiassa tämä tarkoittaa tietyn kausaalisen suhteen eristämistä

muilta: poistamme kaikkien muiden paitsi tarkasteltavan muuttujan vaikutuksen lopputulokseen. Tämä korostaa eristysten instrumentalistista luonnetta. Teemme eristyksiä vain siitä syystä, että voimme niiden avulla arvioida, mikä vaikutus yksittäisen muuttujan manipulaatiolla on lopputulokseen – toisin sanoen, paljastaa kausaalisia suhteita manipulaatioteorian määrittämällä tavalla.

Teoreettiset eristykset ovat tietty tapa toteuttaa interventioita, ja niiden on täten myös täytettävä intervention ehdot. On helppo nähdä, että tämä ei ole ongelma: oletusten avulla tehdyt eristykset täyttävät intervention kriteerit lähes automaattisesti. Mallit ovat kuitenkin tutkijoiden luomia maailmoja, joten heillä on lähes kirjaimellisesti käytössään ”jumalan käsi”, jolla vaikuttaa asioihin oletusten avulla. On syytä muistaa, että intervention kriteerit eivät vaadi intervention toteuttamista käytännössä. Vaikka teoreettiset eristykset eivät määritelmällisesti ole kausaalisia, ne mahdollistavat silti kausaalisten päätelmien tekemisen manipulaatiomallin mukaisesti. Se, ovatko nämä päätelmät sovellettavissa aidossa maailmassa, riippuu siitä, ovatko mallin kuvaamat suhteet toiminnassa aidossa maailmassa, ja kuvaako malli niitä oikein ja uskottavasti. Manipulaatioteoria antaa myös merkityksen sille, miksi voimme olettaa tekijöille käytännössä mahdottomia arvoja: kausaalisen suhteen määritelmään ei sisälly rajoitetta siitä, mihin arvoihin muut muuttujat on kontrolloitava. Olennaista on vain, että on olemassa *joitakin* V:n muiden muuttujien arvoja, joilla *jokin* mahdollisen syyn manipulaatio vaikuttaa lopputulokseen.

Sekä Mäen että Woodwardin teorit ovat syvästi kiinni ajatuksessa testaamisesta. Interventiot voidaan nähdä hyvin rakennettuina kokeina, ja sama pätee malleihin. Aiemmin kuvasimme interventioita ilmapuntariesimerkin avulla, mutta sama tilanne voidaan helposti tulkita myös Mäen viitekehityksessä. Tarkastelun avuksi luomme mallin (kokeen), johon sisältyy myrskyä, ilmanpainetta ja puntarin arvoa kuvaavat muuttujat – tämä on yksinkertaistus todellisuudesta, josta on jätetty valtava määrä asioita pois, ja johon on poimittu vain tutkittavan tapahtuman kannalta relevantit tekijät. Tämän jälkeen suoritamme eristuksen (intervention), jolla eristämme myrskyn ja ilmanpaineen suhteen puntarin arvon vaikutukselta (eli manipuloimme puntarin arvon muuttumattomaksi). Nyt voimme suorittaa ilmanpaineeseen kohdistuvan manipulaation saadaksemme selville,

miten myrskyä kuvaava muuttuja reagoi tähän manipulaatioon. Mallia tarkastelemalla voimme tällä tavoin selvittää sen sisältämien muuttujien välisiä kausaalisuhteita saapuaksemme Woodwardin tyyliiseen selitykseen, joka kertoo, miten myrskyyn voidaan vaikuttaa; tai toisin sanoen, minkä mahdollisen mekanismin kautta myrsky olisi voinut syntyä.

4.2.3 Invarianssi ja uskottavuus

Mallien tulee olla uskottavia sijaissysteemejä. Uskottavuus on olennaista, sillä se mahdollistaa teoreettisten eristysten näkemisen aitoon maailmaan kohdistuvina *mahdollisina* interventioina, mikä puolestaan oikeuttaa mallin sisäisten päätelmien yleistämisen aitoon maailmaan. Mitä uskottavuudella tosiasiasa tarkoitetaan? Tätä kysymystä varten mallit on hyödyllistä nähdä yleistyksinä tiettyjen tekijöiden välisistä suhteista, joiden avulla voidaan vastata m-kysymyksiin niiden sovellusalan sisällä. Kuten aiemmin, sovellusalan voidaan nähdä olevan invarianssin määrittämä. Esitän, että uskottavuuden ja invarianssin käsitteet ovat näin ollen liitettävissä toisiinsa: kysymys siitä, voidaanko jonkin mallin sisäisiä tuloksia käyttää aidossa maailmassa, voidaan ymmärtää kysymyksenä siitä, onko malli riittävän invariantti suhteessa niihin tekijöihin, jotka tapahtumaan aidossa maailmassa vaikuttavat.

Malli ei ole uskottava, jos se ei ole invariantti – Hotellingin mallia ei voida käyttää uskottavana sijaissysteeminä, mikäli sen kuvaamat mekanismit eivät ole invariantteja suhteessa relevantteihin aidon maailman ilmiöihin, kuten kuluttajien epärationaalisuuteen tai maantieteen monimuotoisuuteen. Vastaavasti uskottavuus tarkoittaa invarianssia relevanteilla akseleilla. Hyvien mallien kuvaamat mekanismit eivät siis ole täysin oletustensa rajoittamia: mekanismit ja niiden kuvaukset pysyvät voimassa, vaikka oletuksia löysätään vastaamaan aidon maailman toimintaa. Eli toisin sanoen, tärkeimmät paljastetut mekanismit ovat niitä, jotka pysyvät voimassa, vaikka niitä kuvaaviin malleihin kohdistetaan aidon maailman toimintaa relevantilla tavalla vastaavia interventioita.

Yhteenvedon voidaan todeta, että Mäen ja Woodwardin teorioilla on ainakin kolme keskeistä leikkauspistettä. Ensinnäkin, kausaaliseen selitykseen pyrkivät mallit ovat

tapoja kuvata kausaalirakenteita manipulaatioteoriaa vastaavalla tavalla. Toiseksi, teoreettiset eristykset ovat interventioita, joiden avulla voimme tarkastella ja paljastaa yksittäisiä kausaalisia mekanismeja manipulaatioteorian määritelmien mukaisesti. Kolmantena, mallien käyttökelpoisuus aidossa maailmassa riippuu valitusta muuttujajoukosta ja niiden uskottavuudesta, eli mallin invarianssista suhteessa relevantteihin seikkoihin. Kun sanomme, että mallit ovat selittäviä, tarkoitamme, että ne kuvaavat riittävän todenmukaisesti tapahtuman mahdollista kausaalista rakennetta, ja että niiden avulla voidaan vastata m-kysymyksiin relevantin sovellusalan sisällä.

Siirrytään seuraavaksi tarkastelemaan näitä ajatuksia hieman konkreettisemmin, ja tutustutaan muutamaa erilaiseen taloustieteen selitykseen. Tavoitteenani on osoittaa, että edellinen käsittely soveltuu monien taloustieteen menetelmien ja mallien tulkitsemiseen.

4.3 Esimerkkejä taloustieteestä

4.3.1 Regressioyhtälöt

Regressioyhtälöt ovat keskeinen taloustieteen työkalu, ja niitä on kohdattu myös aiemmin tässä tutkielmassa. Pyritään seuraavaksi aiemman käsittelyn pohjalta selventämään, mitä näiden yhtälöiden tulkitseminen kausaalisesti tarkoittaa.

Tyypillinen monimuuttujaregressiota hyödyntävä malli voisi näyttää seuraavalta:

$$y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + \epsilon_i, \quad i = 1, 2, \dots, n$$

Tässä yhtälössä kuvataan, miten muuttuja Y (josta meillä on havaintoja y_i) riippuu tietyistä selittävistä muuttujista $X_1 \dots X_k$ (joista meillä on havaintoja $x_{1i} \dots x_{ki}$) ja vakiotermit α . Kertoimet $\beta_1 \dots \beta_k$ kuvaavat, miten paljon selitettävä muuttuja y_i muuttuu, kun vastaava selittävä muuttuja $x_{1i} \dots x_{ki}$ muuttuu. Tässä vaiheessa emme oleta tämän yhteyden olevan kausaalinen. Lisäksi malliin kuuluu virhetermi ϵ_i , joka voidaan nähdä joko Y :n mittausvirheenä tai havaintona virhemuuttujasta U , joka sisältää kaikkien mallin ulkopuolelle jätettyjen muuttujien kokonaisvaikutuksen Y :hyn. Tällaiset mallit voidaan helposti tulkita teoreettisten eristysten tuotoksina: poisjättöjen avulla pyrimme nostamaan maailmasta esiin selitettävän tapahtuman Y kannalta keskeiset vaikuttajat.

Mallinnuksen onnistuminen puolestaan riippuu siitä, miten hyvin olemme valinneet muuttujat $X_1 \dots X_k$.

Edeltävän kaltaista regressioyhtälöä voidaan käyttää kuvauksena tietyistä riippuvuussuhteista jonkin datan sisällä, ja sen avulla voidaan pyrkiä tekemään ennusteita perustuen korrelaatioihin selitettävän ja selittävien muuttujien välillä – nämä tavoitteet ovat täysin deskriptiivisiä, eivätkä vaadi minkäänlaista huomiota tapahtuman sisäisistä kausaalisista suhteista. Mikäli haluamme tulkita yhtälön kausaalisesti, se kertoo, miten Y reagoi mahdolliseen X_i :n interventioon. Regressioyhtälö on kausaalinen ja näin ollen selittävä *vain*, jos sen kuvaamaa yhteyttä voidaan käyttää tällä tavoin kuvauksena manipulaatioiden vaikutuksesta.

Mikäli malli tulkitaan kausaalisesti, muuttujat $X_1 \dots X_k$ voidaan nähdä Y :n suorina syinä – lineaarisessa tapauksessa tämä tarkoittaa, että yhden yksikön muutos muuttujassa X_i aiheuttaa β_i :n suuruisen muutoksen Y :ssä. Nämä yhtälön kuvaamat suhteet pätevät kaikille tarkasteltavan populaation jäsenille, ja yksilöllinen vaihtelu riippuu täysin virhetermistä ϵ_i . (Woodward, 2003, s. 321.) On huomioitava lisäksi, että kertoimet $\beta_1 \dots \beta_k$ kuvaavat vastaavan muuttujan $X_1 \dots X_k$ aitoa vaikutusta Y :hyn tutkittavassa populaatiossa. Emme pysty havaitsemaan näitä kertoimia suoraan, vaan meidän on estimoitava ne tutkittavasta otoksesta – meidän on mukautettava mallimme tehtyihin havaintoihin saadaksemme selville valitsemiemme muuttujien vaikutuksen käytettävässä datassa.

Usein paras tapa estimoida regressiokertoimia on pienimmän neliösumman menetelmän avulla. Jotta menetelmän avulla saadut estimaatit ovat harhattomia ja tehokkaita, teemme tiettyjä oletuksia – eli idealisaatioita – koskien mallin virhetermiä. Näiden oletusten (jotka tunnetaan Gauss-Markov oletuksina) mukaan: 1) jokaisen yksittäisen havainnon kohdalla virhetermin odotusarvo on nolla, 2) selittävät muuttujat ja virhemuuttuja ovat riippumattomia, 3) jokaisen havainnon kohdalla virhetermillä on sama varianssi (eli se on homoskedastinen), ja 4) virhetermit eivät saa olla korreloituneita keskenään (Verbeek, 2017, ss. 15–16). Mielestäni nämä oletukset ovat tiiviisti kiinni ajatuksessa mallinnuksen onnistumisesta ja intervention ehtojen täyttymisestä. Ensimmäinen ja toinen oletus

vaativat, että muuttujat on valittu oikein, ja että relevantteja muuttujia ei ole jätetty pois mallista. Lisäksi X_i :n interventio ei saa vaikuttaa Y :hyn mitään mallin ulkopuolelle jätettyä reittiä pitkin. Oletukset kolme ja neljä puolestaan takaavat kuvatusuhteen invarianssin: mikäli virhetermin varianssi muuttuu esimerkiksi X_i :n arvojen kasvaessa, funktion kuvaus ei enää pidä täsmällisesti paikkansa, sillä virhetermin jakauma muuttuu olennaisesti – malli ei ole invariantti suhteessa relevantteihin X_i :n interventioihin. Samoin, mikäli virhetermit ovat korreloituneita keskenään (erityisesti aikasarjojen autokorrelaation tapauksessa, jossa havainnot ovat ajallisesti peräkkäisiä, ja virhetermi on korreloitunut edellisen havainnon virhetermin kanssa), funktion kuvaus muuttuu havaintojen määrän kasvaessa rikkoen funktion invarianssin.

Pienimmän neliösumman menetelmän avulla päädyimme seuraavaan malliin:

$$y_i = a + B_1x_{1i} + B_2x_{2i} + \dots + B_kx_{ki} + u_i, \quad i = 1, 2 \dots n$$

Vakiotermi a , kertoimet $B_1 \dots B_k$ ja virhetermi u_i ovat estimaatteja aiemman mallin havaitsemattomista parametreista. Malli voidaan käytännössä nähdä sijaissysteeminä, jossa Gauss-Markov oletukset pätevät, ja näin ollen sen sisällä kertoimet $B_1 \dots B_k$ todella ovat parhaat mahdolliset estimaatit vastaavien muuttujien kausaalista vaikutuksesta. Siirryttäessä aitoon maailmaan tehtyjen oletusten paikkansapitävyys, kuten myös niiden mukainen kerrointen kausaalinen tulkinta, muuttuvat kyseenalaisiksi. Oletusten ja niitä seuraavien johtopäätösten käyttökelpoisuus (toisin sanoen mallin kontrafaktuaalinen uskottavuus ja invarianssi suhteessa relevantteihin aidon maailman muuttujiin) tulee arvioida tapauskohtaisesti. Yhtälön kuvaaman kausaalisen suhteen olemassaolon kannalta on välttämätöntä, että on olemassa jokin skaala X_i :n interventioita, joiden tapahtuessa funktio tarjoaa riittävän tarkan kuvauksen tätä seuraavasta Y :n muutoksesta – eli toisin sanoen, tehdyn yleistyksen on oltava tietyn asteisesti ja relevantisti invariantti suhteessa näihin interventioihin. Mikäli X_i nähdään Y :n suorana syynä, tämä vaatii intervention ehtojen mukaisesti, että tehty interventio ei muuta tai ole korreloitunut minkään muun selittävän muuttujan kanssa, virhetermi mukaan lukien. (Woodward, 2003, s. 322.) Oletukset ovat vain keino päästä tähän päämäärään – tai toisin

sanoen, keino eristää tutkittava kausaalinen suhde. Tehtävien interventioden ehdot muuttuvat luonnollisesti tarkasteltavan muuttujan syyn tyyppin mukaisesti.

Edeltävän käsittelyn tavoitteena on selventää, mitä tarkoitamme, kun tulkitsemme regressioyhtälöitä kausaalisesti ja minkä ehtojen vallitessa tämä tulkinta on tosi. Lisäksi toivon sen toimineen osoituksena siitä, että *kausaaliset* regressioyhtälöt vastaavat manipulaatioteorian tulkintaa kausaatiosta ja selityksistä, ja että ne voidaan tavoittaa aiemmin kuvatun kaltaisen mallinnuksen keinoin. Kysymys siitä, onko meillä syytä uskoa jonkin vastaavan kausaalisen väitteen totuuteen, on tarkoituksella jätetty auki. Tämä vaatii kausaalista päättelyä, johon tutustutaan lyhyesti seuraavaksi.

4.3.2 Kausaalinen päättely

Tieto kausaatiosta on tietoa kontrolloinnin mahdollisuudesta ja manipulaatiosuhteista. Meillä on lukemattomia syitä olla kiinnostuneita näistä asioista, ja kaikkien politiikkavaikutusten arvioinnissa haluamme tietää, mikä vaikutus *juuri tällä* muutoksella oli tarkasteltavaan seikkaan. Miten ihmisten kokemus hyvinvoinnista muuttuu, jos koulutuksen hintaa lasketaan? Entä mitä tapahtuu tuloille? Suoraviivaisin tapa vastata tähän kausaaliseen kysymykseen on toteuttaa tarkasteltava ehdotus: lasketaan koulutuksen hintaa, ja selvitetään miten tutkittavat muuttajat reagoivat. Tähän liittyy kuitenkin ongelma, sillä kuten Varian (2016, s. 7311) asian hienosti ilmaisee: ”haluamme vastauksen *ceteris paribus* kysymykseen, mutta datamme on luotu *mutatis mutandis*”.⁷ Kun toteutamme laajan uudistuksen, emme pysty pitämään kaikkia muita muuttujia kontrolloituina. Tämä tekee tarkan kausaalisen vaikutuksen arvioinnista mahdotonta. Taloustieteessä on viime vuosina tullut yhä suosittumaksi hyödyntää menetelmiä, jotka kuitenkin mahdollistavat kausaalisten päätelmien tekemisen – ja jotka mukailevat saumattomasti manipulaatioteorian viitekehystä.

Tutkijan työkalupakin tyyppisin (ja kaikkein voimakkain) keino arvioida kausaalista vaikutusta on kontrolloidun satunnaiskokeen avulla. Tässä koeasetelmassa kohdejoukko

⁷ Näkemys, jonka mukaan kausaalinen vaikutus määritellään koetulosten ja kontrollituloksen erotuksena, tunnetaan Rubinin kausaalimallina (ks. Rubin & John, 2011). Kuten Varian asiaa kuvaa, tätä määritelmää käytettäessä joudutaan vastaamaan niin kutsuttuun kausaalisen inferenssin fundamentaaliseen ongelmaan: emme voi millekään yhdelle tutkimuskohteelle havaita sekä kontrolli- että koetulosta.

jaetaan satunnaisesti koe- ja kontrolliryhmiin, ja koeryhmään kohdistetaan jokin käsittely. Käsittelyn kausaalinen vaikutus saadaan seuraavasta yhtälöstä:

$$\begin{aligned} & \text{Vaikutus koeryhmässä} - \text{Vaikutus kontrolliryhmässä} \\ &= [\text{Vaikutus koeryhmässä} - \text{Vaikutus koeryhmässä ilman käsittelyä}] \\ &+ [\text{Vaikutus koeryhmässä ilman käsittelyä} - \text{Vaikutus kontrolliryhmässä}] \\ &= \text{Käsittelyn vaikutus koeryhmässä} + \text{valintaharha. (Varian, 2016, s. 7311.)} \end{aligned}$$

Perusajatus kontrolloidun satunnaiskokeen takana on kontrolliryhmän avulla luoda uskottava kontrafaktuaali tilanteesta, jossa koeryhmään ei kohdisteta käsittelyä. Näin saadaan selville yksittäisen intervention vaikutus koeryhmässä, sillä täysin satunnaisesti valituissa ryhmissä valintaharhan odotusarvo on nolla. Kausaalinen vaikutus saadaan selville vertaamalla kontrafaktuaaleja täysin manipulaatioteorian mukaisesti. Näin täydellinen koeasetelma ei kuitenkaan aina ole mahdollinen, etenkin taloustieteessä. Tästä syystä usein joudutaan turvautumaan kvasikokeellisiin tutkimusasetelmiin, joiden voidaan myös nähdä eri tavoin pyrkivän tavoittamaan manipulaatioteorian vaatimukset onnistuneista interventioista.

Difference-in-differences -asetelmaa voidaan hyödyntää esimerkiksi tutkittaessa verotuksen muutosta. Saadaksemme selville veromuutoksen vaikutuksen johonkin muuttajaan, jaamme populaation satunnaisesti kahteen ryhmään, joista toisen veroja muutetaan ja toisen ei. Tämän jälkeen pyrimme tulkitsemaan syntyneen eron ryhmien välillä toteutetun intervention kausaalisenä vaikutuksena. Jotta tämä päätelmä on oikeutettu, tiettyjen ehtojen on täytyttävä. Ensinnäkin, tehty interventio ei saa olla korreloitunut minkään yhtäaikaisen muutoksen kanssa. Toiseksi, molempien ryhmien on kehityttävä yhdensuuntaisesti yli ajan – jos ryhmien kehityksen välillä on interventiosta riippumattomia eroja, kausaalisen vaikutuksen arviointi on mahdotonta.⁸ Kolmanneksi, ryhmien koostumuksen täytyy pysyä samana vertailun mahdollistamiseksi, eli tutkimuksen kohteet eivät itse saa vaikuttaa siihen, kohdistetaanko heihin interventio vai ei. Näin voisi tapahtua, jos henkilöt pystyisivät itse toiminnallaan vaikuttamaan

⁸ Kehityksen yhdensuuntaisuuden todistaminen vaatii aikasarjadataa molempien ryhmien osalta, mieluiten mahdollisimman pitkältä ajalta ennen intervention toteuttamista.

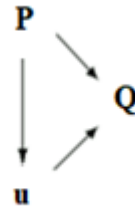
esimerkiksi siihen, siirtyvätkö he veromuutoksen piiriin vai eivät. Viimeisenä, tehty interventio ei saa aiheuttaa kontrolliryhmään kohdistuvia heijastusvaikutuksia. Kuten kontrolloidun satunnaiskokeen kohdalla, tämä koeasetelma pyrkii luomaan manipulaatioteorian mukaisen kontrafaktuaalin, jota vasten intervention vaikutusta voidaan tarkastella. Edellä kuvatut ehdot pyrkivät takaamaan kontrafaktuaalisen tilanteen todenmukaisuuden, jotta tehty intervention vaikutus voidaan tulkita kausaalisesti manipulaatioteorian mukaisesti.

Regression epäjatkuvuusasetelmaa voidaan käyttää tarkastelemaan intervention kausaalista vaikutusta tilanteissa, joissa interventio toteutetaan tarkasteltavan jatkuvan muuttujan jostakin pisteestä eteenpäin. Tavallisin esimerkki koskee stipendien jaon vaikutusta esimerkiksi koulumotivaatioon. Koska stipendit jaetaan tietyn keskiarvon ylittävälle oppilaille, emme voi käyttää difference-in-differences tyyppistä tarkastelua stipendin saaneiden ja muiden oppilaiden välillä, sillä ero ryhmien välillä ei ole satunnainen: hyvin pärjäävät oppilaat ovat jo lähtökohtaisesti motivoituneempia. Ero stipendin saamisen rajan juuri ylittävien ja alittavien oppilaiden välillä voidaan kuitenkin nähdä häviävän pienenä. Jos siis havaitsemme regression muuttuvan epäjatkuvaksi pisteessä, jossa interventio tapahtuu, voimme pitää tätä epäjatkuvuutta mittarina kausaalisesta vaikutuksesta, ja tapahtunut muutos voidaan nähdä määrällisesti intervention kausaalisena vaikutuksena.

Hyvin tyypillinen keino päästä käsiksi kausaaliin päätelmiin on hyödyntää *instrumenttimuuttujia*. Sen ensimmäinen sovellus lienee myös sen tunnetuin esimerkki: vuonna 1928 Philip Wright pyrki arvioimaan regressiomuotoisesti voion tarjonnan joustoa suhteessa voion hintaan. Ikäväkseen Wright sai huomata, että kysyntä- ja tarjontakäyrien muodostaminen tältä pohjalta oli mahdotonta, sillä määrä ja hinta molemmat määräytyvät kysynnän ja tarjonnan perusteella. Toisin sanoen, selittävä muuttuja on korreloitunut virhemuuttujan kanssa, ja tarkkojen estimaattien tekeminen ei ole mahdollista. Wright päätti ratkaista tilanteen keksimällä kolmannen muuttujan, joka vaikuttaa tarjontaan, mutta ei kysyntään. Muuttujaksi valikoitui alueelliset sademäärät: sademäärä vaikuttaa lehmien laiduntamiseen ja samalla tuotettuihin määriin, kuitenkin vaikuttamatta millään tavoin alueelliseen kysyntään. Kun nyt ilmaisemme tuotetun määrän sademäärien kautta,

voimme pitää kysynnän muuttumattomana, ja sademäärän muutoksia tarkastelemalla tutkia, miten tuotettu määrä ja hinta reagoivat muutoksiin tarjonnassa. (Stock & Watson, 2020, ss. 431–432.)

$$\ln(Q_i^{voi}) = \beta_0 + \beta_1 \ln(P_i^{voi}) + u_i$$



Wrightin regressioyhtälö (jossa Q on voion tuotannon määrän ja P voion hinnan havainto i), ja mallin kausaalinen rakenne. Muutokset P:ssä aiheuttavat muutoksia myös virhemuuttujassa, minkä takia P:n yksittäisen vaikutuksen arviointi on mahdotonta. Tämä on vain yksi mahdollinen kuvaus tästä ongelmia tuottavasta suhteesta; voisi myös olla, että kyse on samanaikaisesta kausaatiosta, jolloin P aiheuttaa Q:n ja Q aiheuttaa P:n. Tämäkin tilanne olisi kuitenkin estimoitavissa instrumenttimuuttujien avulla.

Instrumenttia (tässä tapauksessa sateen määrää) merkitään Z:lla, ja siihen perustuva regressio rakennetaan ja estimoidaan kaksivaiheisen pienimmän neliösumman menetelmän avulla seuraavasti:

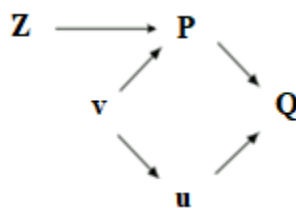
- 1) $P_i^{voi} = \pi_0 + \pi_1 Z_i + v_i$
- 2) Estimoidaan pienimmän neliösumman menetelmän avulla

$$\widehat{P}_i^{voi} = \widehat{\pi}_0 + \widehat{\pi}_1 Z_i$$
- 3) $\ln(Q_i^{voi}) = \widehat{\beta}_0 + \widehat{\beta}_1 \ln(\widehat{P}_i^{voi}) + u_i$

Ensimmäisessä vaiheessa luomme regression, joka jakaa hinnan kahteen osaan: $\pi_0 + \pi_1 Z_i$ on se osa, joka on ennustettavissa Z:n avulla, ja näin ollen se ei ole korreloitunut alkuperäisen virhetermin u_i kanssa, kun taas toinen osa, v_i , kattaa nyt ne hinnan tekijät, jotka ovat korreloituneita alkuperäisen virhetermin kanssa. Toisessa vaiheessa suljemme v_i :n vaikutuksen pois estimoimalla hinta ainoastaan virhetermin kanssa epäkorreloituneiden tekijöiden kautta. Kolmannessa vaiheessa luomme regression toisessa vaiheessa tekemämme estimaatin avulla, ja mikäli instrumentti on hyvin valittu, käytettävän hinnan estimaatin ei tulisi enää olla korreloitunut virhetermin kanssa. Lopulta estimoimme kolmannen vaiheen regressiokertoimet jälleen pienimmän neliösumman

menetelmällä, ja saavumme kaksivaiheisen pienimmän neliösumman estimaatteihin $\widehat{\beta}_0$ ja $\widehat{\beta}_1$. (Stock & Watson, 2020, s. 429.)

Instrumentin on toimiakseen täytettävä kaksi ehtoa: instrumentti ei saa olla korreloitunut alkuperäisen virhemuuttujan kanssa (eksogeenisuus) ja instrumentin on oltava korreloitunut selittävän muuttujan kanssa (relevanssi). Aivan kuten aiemmin regressioyhtälöiden kohdalla, instrumentti voi siis olla ei-kausallinen kuvaus riippuvuussuhteista käytettävässä datassa. Instrumentin kausaalisuudella ei käytännössä ole vaikutusta sen avulla saavutettaviin johtopäätöksiin alkuperäisen selittävän muuttujan kausaalisesta vaikutuksesta. Yksinkertaisuuden vuoksi kuvitellaan instrumentti nyt kausaalisena – tällöin instrumenttimuuttujat ovat käytännön sovellus osittaisen syyn määritelmästä.⁹ Ajatus on, että instrumenttimme vaikuttaa selitettävään muuttujaan ainoastaan valitsemamme selittävän muuttujan kautta – nyt Z:aa manipuloimalla saamme selville, kuinka Q reagoi Z:n aiheuttamiin muutoksiin P:ssä (pitäen kaikki muut reitit kontrolloituina). Jos malli kuvaa todenmukaisesti selitettävää tapahtumaa ja sen kausaalista rakennetta, ja toteutetut interventiot täyttävät näin ollen manipulaatioteorian ehdot, alla kuvatun kaltainen instrumenttiregressio paljastaa yksittäisen P:n intervention kausaalisen vaikutuksen Q:hun pelkästään niiden välisen suoran reitin kautta.



Instrumenttiregression kausaalinen rakenne, jos valittu instrumentti on kausaalinen

Käsiteltyjen esimerkkien kautta olen pyrkinyt osoittamaan, että halutessamme taloustieteessä tehdä kausaalisia päätelmiä, tutkimusasetelmien on eristettävä tutkittava kausaalinen suhde. Kausaalisen vaikutuksen suuruuden arviointi on mahdollista ainoastaan tämän erityksen pitäessä. Kontrolloidut satunnaiskokeet ja kvasikokeelliset

⁹ Mikäli kyseessä on pelkkä korrelaatio, tapahtuman kausaalinen rakenne (ja instrumentin suhde siihen) olisi luonnollisesti erilainen. Erityisesti pienissä otoksissa on mahdollista, että korrelaatio on täysin sattumanvaraista, jolloin tapahtumien välillä ei ole minkäänlaista kausaalista rakennetta – niin kauan kun tämä korrelaatio kuitenkin pätee, instrumentti voi olla käyttökelpoinen, koska tavoitteemme kannalta se käyttäytyy kuin kausaalinen tekijä.

menetelmät pyrkivät aikaansaamaan tämän luomalla kontrafaktuaalin, joka on kaikin tavoin identtinen tutkimustilanteen kanssa, pois lukien tutkijan toteuttama interventio koeryhmään. Samaan pyrkii instrumenttiregressio, ainoastaan jo olemassa olevan datan kohdalla – se pyrkii eristämään tietyn suoran syysuhteen muilta mahdollisilta vaikuttavilta reiteiltä. Huomioitavaa on, että kausaalinen päättely vaatii tästä huolimatta *aina* induktiivisen hypyn, sillä verrattavat tilanteet eivät koskaan voi täysin täyttää *ceteris paribus* vaatimusta. Manipulaatioteorian voidaan kuitenkin nähdä luovan pohjan ja antavan merkityksen näiden tutkimusasetelmien toimivuudelle ja käytölle kausaalisten tulkintojen tekemisessä.

4.3.3 Tasapainoselitykset

Jo tämän työn alussa esitettiin lyhyesti kysymys siitä, ovatko tasapainoselitykset kausaalisia. Kuuluisin vastaesimerkki tästä on peräisin Elliott Soberilta (1983). Sober käytti esimerkkinään sukupuolten jakautumista eri populaatioissa – tämä on esimerkki tilanteesta, jossa systeemi hakeutuu tasapainoon tiettyjen mikrotason mekanismien kautta. Parit, jotka tuottavat enemmän vähemmistösukupuoleen kuuluvia jälkeläisiä, saavat lisääntymisedun merkittävästi 1:1 tasapainosta poikkeavissa populaatioissa, sillä he saavat myös keskimääräisesti enemmän lapsenlapsia. Tämän myötä yleistyvät myös ne perinnölliset tekijät, jotka aiheuttavat suuremman todennäköisyyden tuottaa vähemmistösukupuolta edustavia jälkeläisiä, mikä palauttaa populaation hitaasti takaisin 1:1 tasapainoon. Entä jos nyt havaitsemme tällaisen kehityksen tapahtuneen? Miten selitämme sen, että tietyn populaation naispuolisten jäsenten osuus on laskenut 73 prosentista viiteenkymmeneen? Soberin (1983, s. 207) mukaan *kausaalinen* selitys tälle tapahtumalle on kuvaus populaation mikrotason lisääntymiskäyttäytymisestä. Tällainen selitys kertoo ainoastaan, mihin suuntaan populaatio on kehittymässä, ja sen, miten lopputulokseen päädyttiin. Tähän verrattuna *tasapainoselitys* tavoittaa Soberin mukaan jotain syvempää: se kertoo, miksi sukupuolten välinen ero tasaantui (ja sen oli pitkällä aikavälillä välttämättä tasaannuttava) sitä määrittävien mekanismien syvemmän rakenteen kautta. Tasapainoselitys ei kuitenkaan tee minkäänlaista viittausta niihin mikrotason syihin, jotka tutkittavan ilmiön konkreettisesti tuottavat, ja tasapainoselitykset eivät täten ole kausaalisia.

Woodwardin mallin mukaan Soberin johtopäätös perustuu virheelliseen näkemykseen siitä, mitä kausaalisen selityksen antaminen tarkoittaa. Olemme kohdanneet vastaavan tasapainoselityksen jo monopoliesimerkin kohdalla, ja huomioineet, että tässä tapauksessa selittävyys perustuu nimenomaan yleisten mekanismien ilmaistamiseen, *ei* mikrotason kausaalitapahtumien kuvaukseen. Selitys kertoo, miten selitettävä asia riippuu johdonmukaisesti selityksessä kuvatuista tekijöistä – eli selitys kuvaa ne mekanismit, jotka johtavat tasapainoon. Näitä mekanismeja voidaan käyttää vastaamaan lukuisiin m-kysymyksiin koskien erilaisia sukupuolijakaumia, ja samalla on annettu kausaalinen selitys vallitsevalle tasapainotilanteelle.

Manipulaatioteorian tarkan muotoilun kannalta tasapainoselityksillä on kuitenkin yksi mahdollisesti huolestuttava ominaisuus. Selitykset rakentuvat alku- ja lopputilan välisten mekanististen riippuvuussuhteiden varaan, mutta laajat tasapainomallit ovat käytännössä täysin riippumattomia valitusta alkutilasta. Muistetaan, että kausaalisuhteen olemassaolo vaatii jonkinlaisen invariantin manipulaatiosuhteen olemassaoloa – selittävää muuttujaa manipuloimalla meidän on pystyttävä edes jotenkin saamaan aikaan muutos selitettävässä muuttujassa. Globaaleissa ja täysin vakaisissa tasapainotilanteissa tällaisia interventioita ei määritelmällisestikään ole olemassa, joten näillä tasapainoilla ei tulisi manipulaatioteorian mukaan olla myöskään selitysvoimaa – näin ei selvästikään voi olla, sillä juuri vakaimmat tasapainotilanteet tuntuvat kaikkein selittävimiltä. (Kuorikoski, 2010a, s. 88.)

Kuorikoski pyrkii ratkaisemaan ongelman tarkentamalla tasapainoselitysten explanandumia ja niiden yhteyttä kausaaliselityksiin. Ensinnäkin, tasapainoselitykset *eivät* ole yksittäisten tapahtumien kausaaliselityksiä, vaan ne ovat konstitutiivisia selityksiä systeemitason ominaisuuksista – toisin kuin Sober ajatteli, tasapainoselitysten selityksellinen suhde ei ole alku- ja tasapainotilan välillä. Tasapainoon hakeutuminen on tiettyjen systeemien rakenteellinen ominaisuus, jota hyödyntämällä voimme muun muassa tuottaa kausaalisia selityksiä. (Kuorikoski, 2010a, s. 88.) Tämä ajatus on tiiviisti yhteydessä alaluvussa 1.3 esitettyyn hienovaraiseen eroon konstitutiivisten ja kausaalisten selitysten välillä. Konstitutiivisessa tapauksessa explanandum vaatii kuvauksen niistä mekanismeista, jotka tasapainon aiheuttavat, ja selitys on tällöin vain kuvaus tietystä

systemin ei-kausallisesta rakenteellisesta ominaisuudesta (eli sen taipumuksesta hakeutua tasapainoon). Kausallisessa tapauksessa explanandum koskee jotain yksittäistä selitystilannetta, joka voidaan nyt selittää tämän ominaisuuden avulla. Jos tiedämme jonkin systeemin hakeutuvan tasapainoon, voimme tuottaa hyvin triviaalin selityksen sille, miksi kyseinen systeemi on tasapainossa: systeemillä on ominaisuus olla tasapainossa, jos se ei ole lähiaikoina kokenut shokkeja, ja systeemi ei ole lähiaikoina kokenut shokkeja. Tämän kausallisen selityksen kannalta kausallisen historian kuvaaminen on kuitenkin välttämätöntä. (Kuorikoski, 2010a, ss. 91–92.) Aiemmin kohdattu monopoliesimerkki perustuu samaan ajatukseen: tasapainon avulla luodaan alkutilanne, minkä jälkeen kohdistamme systeemiin shokin, ja pyrimme siihen, että selitettävä tapahtuma on shokin kohdanneen systeemin uusi tasapainotila.

Tasapainoselitykset ovat olennaisia kausalliselityksen kannalta, sillä niiden avulla voidaan suhteellisen yksinkertaisista oletuksista johtaa valtava määrä vastauksia relevantteihin m-kysymyksiin. Tällä tavoin tulkittuna tasapainoselitykset eivät ole kausaalisia selityksiä, mutta niillä on tärkeä rooli monien manipulaatioteorian mukaisten kausaalisten selitysten rakentamisessa, mikä osaltaan selittää niiden laajaa roolia taloustieteellisessä (kausallisessakin) selittämisessä.

5 Pohdinta

Tässä tutkielmassa on pyritty neljään tavoitteeseen: 1) esitellä Woodwardin selitysteoria ja sen vahvuudet keskeisiltä osiltaan, 2) esitellä Mäen teoria taloustieteellisestä mallinnuksesta keskeisiltä osiltaan, 3) kuvata yhteyksiä Woodwardin ja Mäen teorioiden välillä, ja 4) osoittaa, että taloustieteellisiä malleja voidaan tulkita kausaalisesti esitettyjen teorioiden kautta. Käytännössä lopputuloksena on yhdenlainen esitys siitä, mitä kausaatiolla ja kausalliselityksillä voidaan tarkoittaa, ja että miten tätä näkemystä voidaan toteuttaa taloustieteessä. Jotta suhde kahden tapahtuman välillä olisi kausaalinen, sen on oltava invariantti, ja sen avulla tulee voida arvioida interventioita koskevia kontrafaktuaaleja. Selitys puolestaan rakentuu näiden suhteiden varaan siten, että se näyttää, miten tapahtuma seuraa johdonmukaisesti sen kausaalisista syistä, ja tämän kautta, kuinka tapahtumaan pystytään vaikuttamaan. Taloustieteellinen mallinnus

mukailee näitä ajatuksia: Mäen näkemysten mukaisesti mallien rakentamisen metodologia sekä niistä tehdyt kausaaliset tulkinnat perustuvat edellä kuvattuun näkemykseen.

Yksi tutkielman keskeisistä tavoitteista koskee Mäen ja Woodwardin teorian yhdistämistä konseptuaalisesti, ja tämä muodostaa myös työn omintakeisimman osan. Tätä lähestymistapaa motivoi yksinkertainen havainto siitä, että taloustieteelliset mallit tuntuvat mukailevan manipulaatioteorian näkemystä kausaatiosta, ja Mäen näkemys malleista puolestaan vaikuttaa tarjoavan hedelmällisen tavan selittää tätä yhteyttä. Tämän kautta on pyritty luomaan filosofisesti perusteltu, mutta käytäntöä lähellä oleva kokonaiskuva kausaalisesta selittämisestä taloustieteen alalla. Samalla se tarjoaa tuoreen luennan erityisesti Mäen teoriasta, ja kenties auttaa hieman selventämään taloustieteellisen mallinnuksen motivaatiota tässä mielessä. Tältä osin tutkielmalla voidaan nähdä olevan myös kaksi yleisempää tavoitetta: ensinnäkin osoittaa, että manipulaatioteoria on käyttökelpoinen muotoilu monille taloustieteellistä tutkimusta ohjaaville intuitioille – ja että se näin ollen vaikuttaa implisiittisesti monien olemassa olevien taloustieteellisten selitysten ja mallien takana – ja toiseksi, että meidän on lisäksi aiheellista tulkita kausaatiota taloustieteessä tällä tavalla. Nämä tavoitteet on saavutettu siinä, missä tutkielma on pystynyt osoittamaan, että Woodwardin ja Mäen näkemykset muodostavat erittäin käyttökelpoisen kokonaisuuden kausaation ja kausaaliselitysten ymmärtämisessä taloustieteen alalla.

Kausaatioon ja kausaaliselityksiin keskittyminen kattaa suuren osan taloustieteellisistä pyrkimyksistä, mutta ei missään nimessä kaikkea. Erityisesti makrotaloustieteen piiriin kuuluvissa tutkimuskysymyksissä keskitytään usein myös pelkkien ennusteiden tuottamiseen, jolloin tieto korrelaatioista on riittävää, ja selityksiä ei edes tavoitella. Tällaisilla täysin korrelaatioon perustuvilla malleilla on kuitenkin mielenkiintoinen yhteys kausaalimalleihin siinä, että ne käyttäytyvät ikään kuin kausaalimallit kuvaamansa datan sisällä. Pinnallisesti ainoa ero on se, että muuttujien välinen korrelaatio hajoaa aitojen interventoiden edessä, ja toista manipuloimalla ei *oikeasti* voi vaikuttaa toiseen. Käytännössä mallit voidaan kuitenkin esittää samalla tavalla graafeina, jossa tietyt asiat vaikuttavat toisiin (vaikkakin vain näennäisesti), yhden muuttujan kasvu liittyy

olennaisesti toisen muuttujan kasvuun, ja tämä riippuvuussuhde on tietyn asteisesti invariantti kuvaamansa datan sisällä. On siis vain luonnollista, että kausaatio ja kausaalisten päätelmien tekeminen aiheuttaa hämmennystä, erityisesti mallien monimutkaistuesssa merkittävästi. Toivon, että tämä työ osaltaan auttaa jäsentämään tätä eroa, ja tekemään selvemmäksi sen, mitä kausaatiolla tulee tarkoittaa: tieto kausaatiosta on tietoa kontrolloinnin mahdollisuudesta ja interventioiden vaikutuksesta.

Yleisesityksiin pyrkivät projektit joutuvat aina tekemään uhrauksia yksityiskohtien kustannuksella, ja sama pätee myös tämän työn kohdalla. Erityisesti ongelma korostuu, kun keskeisenä lähdemateriaalina on niin poikkeuksellisen tarkasti kirjoitettu teos kuin *Making Things Happen*. Tässä tutkielmassa tarkastelu on rajoitettu koskemaan niitä seikkoja, jotka ovat tärkeitä kausaalisten selitysten kannalta taloustieteen viitekehyksessä – silti tämänkin rajauksen sisällä kattavan esityksen luominen on sisällöllisesti intensiivinen tehtävä. Tästä syystä tämä tutkielma käsittelee hyvin pinnallisesti poikkeustapauksia (joita on luonnollisesti lukuisia), eikä sisällä erityisesti ajankohtaista kritiikkiä tai vastanäkemyksiä esitettyä teoriaa vastaan. Osin tämä voidaan nähdä heikkoutena, mutta toisaalta tämän työn ei niinkään ole tarkoitus esittää kritiikkiä, vaan lähinnä esittää jotain, mitä kritisoida ja jonka päälle tulevaisuudessa rakentaa.

Iso osa niistä kysymyksistä, joita olen joutunut jättämään huomiotta, koskevat esitettyjen näkemysten soveltamista tiettyihin tilanteisiin ja mallityyppeihin. Hedelmällisiä tutkimuskohteita tulevaisuudessa voisivat olla esimerkiksi aggregoidut makromallit (kuten DSGE-mallit), joiden voidaan teoriassa nähdä rakentuvan samanlaisten huomioiden varaan kuin regressio- ja tasapainomallit, mutta joiden yhteys kausaatioon ja mallinnettujen tapahtumien mikrotason kausaali-ilmiöihin ovat vähintäänkin kyseenalaisia (ks. Kuorikoski & Lehtinen, 2018). Tällaiset aukot tarjoavat erinomaisen lähtökohdan tulevalle tarkastelulle teorian sovellusalasta niin filosofian kuin taloustieteen parissa.

Kuten jo työn alussa todettiin, yksi Woodwardin teorian tavoitteista ja lähtökohdista on tieteellistä toimintaa lähellä olevien näkemysten pelastaminen. Uskon, että taloustieteellisen mallintamisen tulkitseminen Woodwardin teorian kautta toisaalta

mahdollistaa kausaalisen tulkinnan antamisen yhä kasvavalle määrälle taloustieteellisiä selityksiä, ja toisaalta mahdollistaa tiettyjen kausaalisten tulkintojen kyseenalaistamisen perustellusti. Vähintäänkin taloustieteellisen tutkimuksen arvioiminen tästä näkökulmasta auttaa selventämään kausaation merkitystä taloustieteilijöille: tarkkojen konseptien avulla meidän on helpompi määrittää, milloin voimme aidosti puhua kausaatiosta ja milloin emme – ja mitä tosiasiallisesti tavoittelemme, kun tavoittelemme selityksiä. Vastaavasti taloustieteellisten esimerkkien kautta voimme tarkentaa filosofista näkemystämme kausaatiosta ja selityksistä, ja ymmärrämme paremmin, mitä teorioitamme käytännössä vaaditaan. Näiden asioiden linkittäminen toisiinsa on varmasti hyödyllistä, ja toivon, että tämä tutkielma voi osaltaan auttaa tuomaan tutkimusaloja relevanteilta osa-alueiltaan lähemmäs toisiaan.

Lähteet

- Björklund, A., & Kjellström, C. (2002). Estimating the return to investments in education: how useful is the standard Mincer equation? *Economics of Education Review*, 21(3), 195–210.
- Bromberger, S. (1966). Why Questions. Teoksessa R. Colodny (Toim.), *Mind and Cosmos: Essays in Contermporary Science and Philosophy* (ss. 86–111). Pittsburgh: University of Pittsburgh Press.
- Cartwright, N. (1983). *How the Laws of Physics Lie*. Oxford: Clarendon Press.
- Cartwright, N. (2009). If no capacities then no credible worlds. *Erkenntnis*, 70(1), 45–58.
- Dretske, F. (2004). Laws of Nature. Teoksessa J. Carroll (Toim.), *Readings on Laws of Nature* (ss. 16–37). Pittsburgh: University of Pittsburgh Press.
- Friedman, M. (1953). The Methodology of Positive Economics. Teoksessa *Essays in Positive Economics* (ss. 3–43). Chicago: University of Chicago Press.
- Hempel, C. G. (1965). Aspects of Scientific Explanation. Teoksessa *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science* (ss. 331–496). New York: The Free Press.
- Hitchcock, C. (2007). Prevention, Preemption, and the Principle of Sufficient Reason. *Philosophical Review*, 116(4), 495–532.
- Hotelling, H. (1929). Stability in Competition. *The Economic Journal*, 39(153), 41–57.
- Kitcher, P. (1989). Explanatory Unification and the Causal Structure of the World. Teoksessa P. Kitcher & W. C. Salmon (Toim.), *Scientific Explanation* (ss. 410–499). Minneapolis: University of Minnesota Press.
- Kuorikoski, J. (2010a). Explaining With Equilibria. Teoksessa *Society by Numbers: Studies on Model-Based Explanations in the Social Sciences* (ss. 81–99). Philosophical Studies from the University of Helsinki.
- Kuorikoski, J. (2010b). *Society by Numbers: Studies on Model-Based Explanations in the Social Sciences*. University of Helsinki, Helsinki.
- Kuorikoski, J., & Lehtinen, A. (2018). Model selection in macroeconomics: DSGE and ad hocness. *Journal of Economic Methodology*, 25(3), 252–264.
- Mäki, U. (1992). On the Method of Isolation in Economics. Teoksessa C. Dilworth

- (Toim.), *Idealization IV: Intelligibility in Science* (Poznan Stu, ss. 319–354). Amsterdam - Atlanta: Rodopi.
- Mäki, U. (2005). Models are experiments, experiments are models. *Journal of Economic Methodology*, 12(2), 303–315.
- Mäki, U. (2009a). MISSing the World. Models as Isolations and Credible Surrogate Systems. *Erkenntnis*, 70(1), 29–43.
- Mäki, U. (2009b). Realistic Realism about Unrealistic Models. Teoksessa H. Kincaid & D. Ross (Toim.), *The Oxford Handbook of Philosophy of Economics* (ss. 68–98). New York: Oxford University Press.
- Niiniluoto, I. (1983). *Tieteellinen päättely ja selittäminen*. Helsinki: Otava.
- Psillos, S. (2002). *Causation and Explanation*. Stocksfield: Acumen Publishing.
- Reiss, J. (2012). The Explanation Paradox. *Journal of Economic Methodology*, 19(1), 43–62.
- Rodrik, D. (2015). *Economics Rules: Why Economics Works, When It Fails, and How to Tell the Difference*. Oxford: Oxford University Press.
- Rubin, D., & John, L. (2011). Rubin Causal Model. Teoksessa M. Lovric (Toim.), *International Encyclopedia of Statistical Science* (ss. 1263–1265). Berlin, Heidelberg: Springer.
- Russell, B. (1913). On the Notion of Cause. *Proceedings of the Aristotelian Society*, 13(New Series), 1–26.
- Salmon, W. C. (1989). *Four Decades of Scientific Explanation*. Minneapolis: University of Minnesota Press.
- Salmon, W. C. (1998). The Importance of Scientific Understanding. Teoksessa *Causality and Explanation* (ss. 79–91). Oxford: Oxford University Press.
- Scriven, M. (1962). Explanations, Predictions, and Laws. Teoksessa H. Feigl & G. Maxwell (Toim.), *Minnesota Studies in the Philosophy of Science, Vol. 3: Scientific Explanation, Space, and Time* (ss. 170–230). Minneapolis: University of Minnesota Press.
- Skow, B. (2014). Are There Non-Causal Explanations (of Particular Events)? *The British Journal for the Philosophy of Science*, 65(3), 445–467.
- Sober, E. (1983). Equilibrium Explanation. *Philosophical Studies*, 43(2), 201–210.

- Stock, J. H., & Watson, M. W. (2020). *Introduction to Econometrics* (4. p.). Harlow: Pearson Education.
- Strevens, M. (2007). Review of Woodward, *Making Things Happen*. *Philosophy and Phenomenological Research*, 74(1), 233–249.
- Sugden, R. (2000). Credible Worlds: The Status of Theoretical Models in Economics. *Journal of Economic Methodology*, 7(1), 1–31.
- Suppes, P. (1962). Models of Data. *Logic, Methodology, and Philosophy of Science: Proceedings of the 1960 International Congress*, 252–261.
- Varian, H. (2016). Causal inference in economics and marketing. *PNAS*, 113(27), 7310–7315.
- Verbeek, M. (2017). *A Guide to Modern Econometrics* (5. p.). Hoboken: Wiley.
- von Wright, G. H. (1973). Explanation and Understanding. *Teoksessa International Philosophical Quarterly* (Vsk. 13). <https://doi.org/10.5840/ipq197313340>
- Woodward, J. (2003). *Making Things Happen: a Theory of Causal Explanation*. Oxford: Oxford University Press.
- Woodward, J. (2017). *Scientific Explanation* (E. N. Zalta, Toim.). Noudettu osoitteesta <https://plato.stanford.edu/archives/fall2017/entries/scientific-explanation>
- Ylikoski, P. (2013). Causal and Constitutive Explanation Compared. *Erkenntnis*, 78(2), 277–297.
- Ylikoski, P., & Kuorikoski, J. (2010). Dissecting Explanatory Power. *Philosophical Studies*, 148(2), 201–219.