

Jyväskylän yliopisto
Matemaattis-luonnontieteellinen tiedekunta

Työllistääkö aktivointi?

Vaikuttavuusanalyysi havainnoivassa tutkimuksessa

Elina Kokkonen
tilastotieteen pro gradu –tutkielma
31. elokuuta 2007
Tilastotieteen yksikkö

Tiivistelmä

Elina Kokkonen: Työllistääkö aktivointi? Vaikuttavuusanalyysi havainnoivassa tutkimuksessa Tilastotieteen pro gradu –tutkielma, Jyväskylän yliopisto, 31.8.2007.
Sivuja 57, liitteitä 3.

Tutkielman tavoitteena oli selvittää mahdollisuuksia huomioida valikoitumisharha havainnoivassa tutkimusasetelmassa ja tutkia aktivointisuunnitelmien laadinnan vaikuttavuutta työllistymiseen. Vaikuttavuutta tutkittiin tässä työssä epidemiologisen kaltaistetun vertailuasetelman avulla. Vertailuryhmät kaltaistettiin propensiteettipistemäärää käyttäen. Aineisto oli otos työministeriön URA –tietokannasta Etelä-Savon, Lapin, Pohjanmaan ja Uudenmaan TE -keskusalueilta.

Aktivointisuunnitelmia laaditaan pitkäaikaistyöttömille. Niiden työllistämistavoitteiden toteutumista tutkittiin vertailemalla koeryhmää jota oli aktivoitu vuonna 2003 kaltaistettuun kontrolliryhmään jolle ei ollut laadittu aktivointisuunnitelmaa lainkaan. Työllistymistä seurattiin vuosina 2004 ja 2005. Kaltaistusperusteena käytettiin propensiteettipistemäärää, eli ennustettua todennäköisyyttä kuulua koeryhmään. Pistemäärää määritettäessä kaltaistusmuuttujina olivat ikä, sukupuoli, koulutus, TE –keskusalue, vajaakuntoisuus, työnhakuala, aikaisempi osallistuminen kurssille tai sijoitus työvoimapolitiittisena toimenpiteenä, sekä työssäolopäivien lukumäärä vuosina 2001 ja 2002. Kaltaistuksella saatiin vähennettyä taustamuuttujien epätasapainoa havainnoivassa tutkimusasetelmassa. Tuloksena saatiin, ettei aktivointisuunnitelmien laadinta paranna työllistymistä tutkitulla aikavälillä.

Aktivoidut olivat hyvin voimakkaasti valikoitunut joukko pitkäaikaistyöttömiä, joten Rosenbaumin (1995) sensitiivisyysanalyysillä selvitettiin tulosten herkkyyttä mahdollisten havaitsemattomien tekijöiden vaikutuksille. Analyysi perustui käsittelytodennäköisyyksien vedonlyöntisuhteille. Sovellettu sensitiivisyysanalyysi osoitti että tulokset olivat herkkiä ja siihen saattaisivat vaikuttaa työllisyyttä estävät tekijät. Näin ollen kaltaistuksen oletus, jonka mukaan propensiteettipistemäärää laskettaessa käytettyjen taustamuuttujien avulla saatiin huomioitua kaikki valikoitumiseen vaikuttavat seikat riittävän hyvin, ei välttämättä toteudu. Työllistyminen on niin monitahoinen tutkittava, että siihen liittyvien vaikutussuhteiden luotettavaan selvittämisen tarvittaisiin aineisto jossa olisi mukana myös esimerkiksi tässä huomiotta jääneitä epävirallisia sosiaalisia suhteita ja työllistymismotivaatiota kuvaavia muuttujia.

Avainsanoja: valikoitumisharha, havainnoiva tutkimusasetelma, vaikuttavuus, propensiteettipistemäärä, kaltaistus, käsittelytodennäköisyys, sensitiivisyysanalyysi

Sisällysluettelo

1. Alkusanat	1
2. Taustaa: Työttömien aktivointi	1
2.1 Rakenteistunut työttömyys ja aktiivinen työvoimapolitiikka	1
2.2 Laki kuntouttavasta työtoiminnasta	3
2.3 Aikaisempia tutkimuksia ja tutkielman tavoitteet	4
3. Vaikuttavuus ja satunnaisuus	7
3.1 Vaikuttavuuden tutkiminen	7
3.2 Satunnaistettu koe: ideaalitalanne	9
3.3 Käsittelyvaikutus ja testaus satunnaistetussa koetilanteessa	11
3.3.1 Käsittelyvaikutus	11
3.3.2 Summatestit ja McNemarin testisuure	12
4. Harha ja havainnoiva tutkimus	15
4.1 Havainnoivan tutkimuksen ero satunnaistettuun koetilanteeseen	15
4.2 Harha	16
4.3 Havaitun valikoitumisharhan kontrollointi	17
4.3.1 Ositus	17
4.3.2 Kaltaistus	18
4.3.3 Propensiteettipistemäärä	19
4.4 Sensitiivisyysanalyysi ja piilevän harhan vaikutuksen estimointi	21
4.4.1 Kokeen sensitiivisyys	22
4.4.2 Merkitsevyytasojen sensitiivisyys	25
5. Sovellus: Aktivoiko aktivointi?	28
5.1 Aineisto ja asetelma	28
5.1.1 URA -rekisteri	28
5.1.2 Kohde, rajaus ja työllistymisen määrittely	29
5.1.3 Otanta	31
5.1.4 Asetelma	32

5.2	Aineiston kuvailua	34
5.2.1	Muuttujat	34
5.2.2	Alustavat tarkastelut	37
5.3	Harhan kontrollointi ja käsittelyvaikutus	40
5.3.1	Harha	40
5.3.2	Kaltaistus	41
5.3.3	Käsittelyvaikutuksen testaus	44
5.3.4	Sensitiivisyysanalyysi	46
5.4	Yhteenveto ja tulosten tulkinta	48
6.	Johtopäätökset	50
	<i>Lähdeluettelo</i>	<i>51</i>
	<i>LIITE 1. Riski ja uskottavuuspäätely</i>	<i>54</i>
	<i>LIITE 2: Muuttujat ja työllistyminen aktiivisuuden suhteen vuonna 2004</i>	<i>56</i>
	<i>LIITE 3: Ennustettujen todennäköisyyksien jakauma</i>	<i>57</i>

1. Alkusanat

Vaikeasti työllistyvien määrä on Suomessa edelleen suuri, vaikka aktiivisen työvoimapolitiikan mukaisesti työllistymistä on pyritty helpottamaan erilaisilla toimilla. Vuonna 2001 tuli voimaan laki kuntouttavasta työtoiminnasta (189/2001), ja lain mukaisesti pitkäaikaistyöttömille on laadittu aktivointisuunnitelmia, ja heillä on mahdollisuus myös kuntouttavaan työtoimintaan.

Tämän työn tarkoituksena on selvittää aktivointisuunnitelmien laadinnan vaikuttavuutta työllistymiseen epidemiologisen kaltaistetun vertailuasetelman ja Rosenbaumin (1995) sensitiivisyysanalyysin avulla. Aktivointi- ja verrokkiryhmät kaltaistettiin kovariaattien suhteen propensiteettipistemäärää käyttäen. Näin pyrittiin vähentämään taustamuuttujien epätasapainoa havainnoivassa tutkimusasetelmassa. Sensitiivisyysanalyysillä selvitettiin tulosten herkkyyttä havaitsemattomille tekijöille.

2. Taustaa: Työttömien aktivointi

2.1 Rakenteistunut työttömyys ja aktiivinen työvoimapolitiikka

Työttömiä ovat henkilöt, jotka eivät ole työsuhteessa, ja jotka ovat ilmoittautuneet työnhakijoiksi. Viime vuosina työttömien kokonaismäärä on vähitellen laskenut, ja vuoden 2007 alkupuolella (tammi-toukokuu) työttömien määrä kuukausittaisella keskiarvolla mitattuna laski noin 202000:een (Työministeriö). Työllisyyden kasvusta huolimatta vaikeasti työllistyvien määrä on asettunut noin 160 000–170 000 henkilön tasolle (Tupo 2-työryhmän mietintö 2006: 28). Työmarkkinoiden pahimpia ongelmia ovat rakenteellinen työttömyys sekä suuret alueelliset työllisyserot: olemme siirtyneet tilanteesta, jossa lama oli työttömyyden pääasiallinen selitys, tilanteeseen, jossa ongelmana on työmarkkinoiden epätasapaino (Soininvaara 1999). Suomessa onkin tilanne jossa samanaikaisesti kun työvoimasta pulaa, on työttömien määrä myös suuri, sillä työpaikkojen täyttöaika on pidentynyt ja kohtaanto ei ole aina hyvä (Tupo 2 - työryhmän mietintö 2006: 7).

Työllisyyspolitiikan puitteissa on käytetty kärkkäitä puheenvuoroja. Suomessa on kuitenkin todellisuudessa toteutettu merkittäviä työllisyys- ja työvoimapolitiittisia uudistuksia, jolla on pyritty pureutumaan työttömyysongelmaan. Aktiivinen työvoimapolitiikka on noussut esiin vaihtoehtona passiiviselle, ainoastaan tukiin perustuvalle politiikalle, mutta hankkeiden tulokset eivät ole yltäneet tavoitteiden tasolle eivätkä toimet ole kiistattomasti edistäneet työllistymistä (Lennon 2005, Tupo 2 - työryhmän mietintö 2006: 8, 10, 28, 32). Aktivoinnin tavoitteista on myös ollut erimielisyyttä: onko tavoitteena henkilön piristäminen ja luoda heille avaimia hallita omaa elämäänsä, vai työttömien itsensä korostama ja lakiin kirjattu työllistyminen (Ala-Kauhaluoma ym. 2004).

Suomessa aktiivisen työvoimapolitiikan keinoja on otettu käyttöön 1990-luvulta lähtien. Aktiivinen työvoimapolitiikka sisältää työnvälitys- ja neuvontapalveluita, sekä työvoimakoulutuksen ja työllistämistoimet (Luukkonen ym. 2005: Ilmakunnas ym. 2001). Aluksi toimet olivat vapaaehtoisia, mutta Suomessa ollaan siirtymässä nyt yhä velvoittavampaan suuntaan. Tupo 2 - työryhmän ehdotusten pohjalta pyritään vuosina 2007–2008 nostamaan aktivointiastetta 2,5 prosenttiyksikköä, joka tarkoittaa aktiiviohjelmiin osallistujien lisäämistä kymmenellä tuhannella. Aktivointiaste on aktivoitujen suhde kaikkiin aktivointiin oikeutettuihin, ja luku vaihtelee suuresti alueittain. Eri toimien tarkoituksena on nostaa työttömien aktivointiaste edellisessä hallitusohjelmassa päätettyyn noin 30 prosenttiin (Tupo 2 - työryhmän mietintö 2006).

Koko palveluprosessin päämäärä on löytää työnhakijalle työtä tai koulutusta. Yhteistyössä laaditut suunnitelmat (eli työllistymisohjelma, työnhakusuunnitelma, yksilöity työnhakusuunnitelma, uudistettu yksilöity työnhakusuunnitelma, aktivointisuunnitelma ja kotoutumissuunnitelma) ovat tärkeä osa tätä palvelua (Tupo 2 - työryhmän mietintö 2006: 14, 18). Suunnitelmat toimivat myös molemminpuolisena sitoumuksena toisaalta työvoimatoimistolle tarjota palveluja, toisaalta työttömälle hakea työtä. Jos henkilö kieltäytyy laatimasta suunnitelmaa, tai osallistumasta sovittuihin toimenpiteisiin, saattaa se vaikuttaa oikeuteen saada passiivitukia.

Vanhasen ensimmäisen hallituksen aikana vaikeasti työllistyvää pitkäaikaistyöttömien joukkoa saatiin kavennettua ja rakenteellinen työttömyys saatiin laskuun vuonna 2004 alkaneen talouskasvun vauhdittamina. Syynä tähän on esitetty aktiivitoimien parempaa vaikuttavuutta. Vaikeasti työllistyvien määrä on silti edelleen suuri, ja työllisyyden nousun odotetaan seuraavina vuosina hidastuvan (Työllisyysohjelman loppuraportti 3/2007: 5, 8, 11). Vanhasen toinen hallitus on asettanut alkaneelle kaudelle kuitenkin yhtä kunnianhimoiset työllisyystavoitteet.

2.2 Laki kuntouttavasta työtoiminnasta

Laki kuntouttavasta työtoiminnasta (189/2001) on sosiaalipoliittinen yritys vaikuttaa työllisyyden kehitykseen. Se sisältää veloitteen laatia pitkäaikaistyöttömälle aktivointisuunnitelma, joka voi sisältää myös kuntouttavaa työtoimintaa. Laki velvoittaa laatimaan aktivointisuunnitelman, kun henkilö on ollut pitkään työttömänä. Työttömille on määritelty ehdot, jotka täytettyään he ovat aktivointiin oikeutettuja: työttömän on tullut saada pitkään työmarkkinatukea tai toimeentulotukea ja hänelle on tullut tehdä työnhakusuunnitelma (tai työnhakuhaastattelu). Alle 25 –vuotiaille suunnitelma laaditaan pikemmin. He ovat aktivointiin oikeutettuja, jos he esimerkiksi ovat saaneet pääasiallisena toimeentulonaan työttömyyden johdosta maksettavaa toimeentulotukea neljä kuukautta, kun 25 vuotta täyttäneiden on täytynyt saada toimeentulotukea 12 kuukautta. Kuntouttava työtoiminta luotiin viimesijaiseksi aktiivisen toiminnan muodoksi.

Tässä tutkielmassa aktivoinnista puhuttaessa tarkoitetaan aktivointisuunnitelman laadintaa. Aktivointisuunnitelma tehdään yhteistyössä työttömän, sosiaalityöntekijän ja työvoimaviranomaisen kanssa. Sen tarkoituksena on yhdessä eri tahojen kanssa edistää henkilön pääsyä avoimille työmarkkinoille. Suunnitelma voi sisältää työllistymistä edistäviä, sosiaali-, terveys-, koulutus- ja kuntoutuspalveluja, sekä kuntouttavaa työtoimintaa (Kuntouttavan työtoiminnan ohjausryhmä 2004: 7).

Kunta päättää kuntouttavan työtoiminnan järjestämisestä. Kuntouttavan työtoiminnan aloitti esimerkiksi vuonna 2004 hieman yli 6 900 henkilöä, joista vajaa 600 oli alle 25-vuotiaita nuoria (Tupo 2 - työryhmän mietintö 2006: 29). Se voi kestää 3-24 kuukautta, mutta yleensä jaksot kestävät keskimäärin vain 2-6 kuukautta ja usein työttömillä on Suomessa työttömyytensä aikana useita erilaisia aktivointijaksoja (Aho ja Kunttu 2001: 7). Jakson jälkeen arvioidaan onko työtoiminta edistänyt valmiuksia palata työhön (Kuntouttavan työtoiminnan ohjausryhmä 2004).

Aktivointisuunnitelma oli tarkoitus tehdä jokaiselle siihen oikeutetulle. Tämä ei kuitenkaan ole ollut mahdollista, sillä toteutus lähti liikkeelle hitaasti. Suunnitelmaan tekoon oikeutettujen henkilöiden määrä nousi vuoden 2004 alkupuolella jo yli 145 000:een ja samaan aikaan aktivointisuunnitelmia oli tehty vasta 43 prosentille kaikista oikeutetuista. Lukumäärälliset tavoitteet jäivät näin ollen saavuttamatta. Työttömät ovat myös eriarvoisessa asemassa eri kunnissa, sillä käytännöt ja lain toimeenpano vaihtelevat niissä suuresti. (Kuntouttavan työtoiminnan ohjausryhmä 2004: 9-11)

2.3 Aikaisempia tutkimuksia ja tutkielman tavoitteet

Työttömyyttä on tutkittu Suomessa laajalti, ja esimerkiksi työministeriön Työpoliittinen tutkimus -sarjassa on ilmestynyt jo 295 nidettä vuoteen 2006 mennessä. Tämän tutkielman kannalta olennaisia menetelmiä ja tuloksia ovat tutkineet eri tahoilta mm. Aho ja Kunttu (2001), Aho ja Koponen (2007), Ala-Kauhaluoma ym. (2004), Hämäläinen (1998), Hämäläinen ja Tuomala (2006) sekä Hämäläinen ym. (2007).

Simo Aho ja Susanna Kunttu (2001) ovat kehittäneet menetelmiä tutkia työvoimapolitiittisten toimien vaikuttavuutta rekisteriaineistojen avulla, joilla voisi selvittää edistävätkö toimet työllistymistä avoimille työmarkkinoille. Lähtökohdaksi he olivat valinneet tietynä ajankohtana työttömänä olleet. Tästä joukosta poimittiin pian sen jälkeen työllistämistoimeen osallistuneet ja näiden kanssa mahdollisimman yhdenmukaiset vertailuryhmät, jotka eivät osallistuneet toimiin. Työllistymistä selvitettiin erilaisilla vaikuttavuusluvuilla. Tutkimuksessa vertailtiin ryhmiä työssäkäynnin määrän keskiarvon (kuukausissa) perusteella. Tutkimuksessa vertailtiin poikkileikkaustietoa ja yhden vuoden seuranta-aikaa, sekä yhden vuoden ja kolmen vuoden seuranta-aikojen perusteella saatuja tuloksia. Olennaista oli että tulokset suhteessa vaikuttavuuteen eivät erityisemmin muuttuneet eri työllistymiskriteereitä käyttäen. (Aho ja Kunttu 2001: 1, 13, 17)

Simo Aho ja Hannu Koponen (2007) seurasivat työvoimapolitiittisiin toimenpiteisiin osallistuneiden työllistymistä vuosina 2002-2005 rekisteriaineistojen avulla. He pyrkivät selvittämään kuinka luotettavaa on työministeriön tavanomainen työllistymisen seuranta, sillä työllistymisestä ei kerry rekistereihin luotettavaa tietoa. Yleensä työllistymistä seurataan kun työttömien henkilöiden toimenpiteisiin osallistumisesta on kulunut kolme kuukautta. Tutkimuksessa todetaan, että seuranta-aikaa pidennettäessä työllistyneiden osuus kasvoi huomattavasti. Taustamuuttujista työllistymisen suhteen eroja löytyi tutkimuksessa iän, koulutustason, vajaakuntoisuuden, työmarkkinahistorian sekä TE -keskusten suhteen. Sukupuoli ei näyttänyt erottelvan työnhakijoita niin selvästi. Aho ja Koponen lisäävät myös, että aktiivitoimenpiteet eivät muuta työllistymistä selittävien taustaseikkojen aiheuttamia eroja työllistymisessä. Erittäin vaikeasti työllistyvissä ryhmissä aktivoinnilla on positiivinen vaikutus työllistymiseen, mutta se ei riitä nostamaan tämän ryhmän työllistymistä alun alkaen paremmat työllistymisedellytykset omaavien tasolle (ks. Aho ja Kunttu 2001). (Aho ja Koponen 2007: 2, 15-18)

Stakesin AKKU -tutkimuksessa (Ala-Kauhaluoma ym. 2004) on käsitelty työttömien aktivointia kuntouttava työtoiminnan lain näkökulmasta. Lain vaikuttavuutta työllistymiseen tarkasteltiin seitsemällä pilotti paikkakunnalla teoreettisen kirjallisuuden, haastattelujen ja havainnointiaineistojen avulla, sekä koko Suomen alueelta 51 kunnasta kyselylomakkein. Tuloksena puolen vuoden seurantajakson aikana aktivoiduista 50 prosentilla työttömyys oli päättynyt, kun kontrolliryhmässä se oli päättynyt 35 prosentilla. Avoimille työmarkkinoille työllistymisessä ei ryhmien välille löydetty eroa. Tutkimusjoukon aktivoiduista 8 prosenttia työllistyi avoimille työmarkkinoille. Sosiaalipoliittisesti tutkimuksen yksi tulos oli, että aktivointisuunnitelmat lisäsivät henkilöiden elämänlaatua, kun henkilöt myös osallistuivat toimenpiteisiin. Asiakkaiden odotukset taas olivat usein korkeat aktivointisuunnitelmien työllistävästä vaikutuksista. Usein lain virallisten tavoitteiden ja annettujen keinojen koettiin olevan ristiriidassa siten, että erilaisia keinoja ei ollut riittävästi. Aktivoinnin tarjoajille ristiriita täsmentyi resursseihin ja tarpeisiin.

Kari Hämäläinen (1998) on selvittänyt aktiivista työvoimapolitiikkaa Suomen työmarkkinoilla taloustieteellisestä näkökulmasta. Hän erittelee monitahoisia taloudellisia seurauksia eri toimenpiteistä Suomessa, todeten aktiivisen työvoimapolitiikan saavutusten olleen vaatimattomia. Hän kritisoi toimenpiteiden vaikuttavuuden arviointia työllistymisellä, sillä työllistymiseen vaikuttavat yksilötason henkilökohtaiset ominaisuudet ja käsittelyryhmiin valikoituminen ei ole satunnaista, vaan riippuu myös mm. yksilöstä itsestään. On erotettava yksilön ominaisuudet intervention vaikutuksista, muuten tulokset ovat harhaisia. Hän tuokin esiin tarpeen tutkia aktiivisten toimenpiteiden vaikutusta työllistymistodennäköisyyteen valikoitumisharhan huomioivilla menetelmillä. Tähän puutteeseen Kari Hämäläinen itse ja Juha Tuomala (2006) vastasivat laajaa mielenkiintoa herättäneessä tutkimuksessaan ”Työvoimapolitiittisten toimenpiteiden vaikutusten arviointi” (Työpoliittinen tutkimus 315). Tutkimus on ei-kokeellinen, ja kohteena ovat erilaiset työvoimapolitiittiset toimenpiteet ja niiden vaikutukset työllistymiseen kaltaistetussa asetelmassa. He käyttivät menetelmiä havaitsemattomien tekijöiden vaikutuksen hahmottamiseksi. Tehokkaimmin työllistäviksi toimiksi osoittautuivat ammatillinen työvoimakoulutus, oppisopimuskoulutus ja yksityiselle sektorille tapahtuva tukityöllistäminen.

Hämäläinen ym. (2007) jatkoivat havainnoivien tutkimusten systemaattisen harhaisuuden arviointia. He vertasivat kahta aiempaa satunnaistettua koetta työnhakukoulutuksen vaikuttavuudesta ei-kokeellisen asetelman tuloksiin. Tuloksena he toteavat, että tavallisesti toimenpiteiden arviointiin käytetyt ei-kokeelliset menetelmät yliarvioivat niiden vaikutuksen työllistymiseen.

Tutkielman tavoitteet

Tämän tutkielman tavoitteena on selvittää aktivointisuunnitelmien laadinnan vaikutuksia työllistymiseen. Tavoitteet on tiivistetty kolmeen kohtaan:

1. Esitellä tapoja huomioida valikoitumisharha havainnoivassa tutkimusasetelmassa.
2. Tutkia kuinka suuri osa aktivoituista työllistyi avoimille työmarkkinoille verrattuna ei-aktivoituihin eli toteutuivatko lain työllistämistavoitteet.
3. Tutkia kuinka herkkä saatu tulos on havaitsemattomien tekijöiden vaikutuksille.

3. Vaikuttavuus ja satunnaisuus

3.1 Vaikuttavuuden tutkiminen

Tilastotieteessä intervention vaikutusten tutkiminen liitetään usein kausaalipäätelyyn. Tilastotieteellisen kausaalipäätelyn tunnetuimpia tutkijoita ovat Donald B. Rubin (mm. 1977), Paul R. Rosenbaum (mm. 1995), David R. Cox (mm. 1972), Paul Holland (mm. 1986) sekä Judea Pearl (mm.1997). Tässä työssä rajoitutaan Rubinin ja Rosenbaumin lähestymistapoihin havainnoivassa tutkimuksessa erityisesti virhelähteiden hallinnan osalta. Niin kutsuttu Rubinin malli perustuu satunnaistettuihin asetelmiin, mutta hän on myös julkaissut artikkeleita tähän tutkielmaan liittyen propensiteettipistemäärän käytöstä havainnoivissa tutkimusasetelmissä. Rosenbaum on keskittynyt havainnoivan tutkimuksen virhelähteiden kontrollointia käsitteleviin menetelmiin, ja on kehittänyt menetelmiä jotka huomioivat valikoitumisharhan vaikutuksen.

Intervention evaluaatiolla tarkoitetaan eritoten terveydenhuollon toimenpiteiden tehokkuutta ja niiden arviointia (Beagerhole ym. 1993: 3-4). Näitä menetelmiä voidaan käyttää myös sosiaalipoliittisten toimien tehokkuuden arviointiin. Vaikuttavuuden tutkimuksessa pyritään löytämään kausaalisuhte oletetun syyn ja seurauksen välille, esimerkiksi mikä on toimenpiteen vaikutus työllistymiseen (tietyissä tilanteessa) verrattuna siihen ettei toimenpidettä tehdä. Kausaalisuhdetta ei voi suoraan nähdä, joten on valittava kausaalimalli tilanteen mukaan. Kausaalinäkemys on tässä sidottu peräkkäisyyteen (Humelainen kausaalisuus) ja todennäköisyyksiin (probabilistinen kausaalisuus). Merkitään nyt T :lla syytä, esimerkiksi käsittelyä, ja Y :llä seurausta eli vastetta. Tällöin seuraavien ehtojen on toteuduttava, jotta voidaan puhua kausaalisuudesta:

- a.) T ja Y on esiinnyttävä yhdessä: jos T , niin todennäköisesti Y .
- b.) T :n on esiinnyttävä ajallisesti ennen Y :tä.
- c.) On oltava olemassa teoreettisesti perusteltu syy, miksi Y aiheutuu T :stä.
- d.) Empiirinen osoitus T :n ja Y :n yhteydestä.

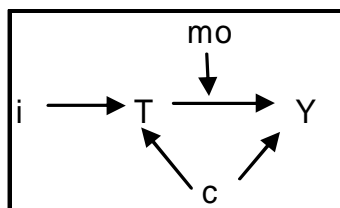
(Dahler-Larsen ja Krogstrup 2003: 11; C. S. Peirce 1868: 90–108; Seidenfeld 1979: 72–73 mukaan)

Jotta yhteys voidaan osoittaa empiirisesti, on oltava tietoa taustamuuttujista x , jotka eivät ole seurausta syy-muuttujasta. Kausaalipäätelyssä oletetaan, että tutkijalla on aineistossa tarpeellinen määrä tietoa näissä havaituissa muuttujissa x , jotta syy-seuraussuhde voidaan jollain tapaa todentaa. Tilastotieteen kannalta on olennaista huomioida miten yksilöt valikoituvat käsittelyyn. Merkitään nyt muuttujalla Z toimenpiteeseen valikoitumista. Nyt Z_i on satunnaismuuttuja, joka kertoo mihin käsittelyryhmään yksilö i kuuluu. Jos toteutunut käsittelyryhmä on z_i yksilölle i , niin yksilön i käsittelyyn valikoitumisen todennäköisyys on

$$p_i = P(Z_i = z_i).$$

Valikoitumistodennäköisyyttä, tai riskiä, voidaan arvioida tapausten määrällä suhteessa kaikkiin mahdollisuuksiin. Tähän päästään uskottavuuspäätelyn kautta (Liite 1). Satunnaistetuissa koeasetelmissa tutkittavat yksilöt valikoituvat satunnaistamisen kautta käsittelyyn T . Useat yhteiskunnalliset asetelmat ovat kuitenkin havainnoivia tutkimuksia, jolloin valikoitumisen Z mekanismi voi vaikuttaa tuloksiin.

Vaikuttavuutta voidaan arvioida jakamalla syy- ja seuraustekijöihin vaikuttavat tekijät väliin tuleviin, sekoittaviin ja indikaattoreihin. Varianssilähtöisessä strategiassa todetaan kausaalisuhte luomalla looginen yhteys kontrolloimalla kaikkien muiden vaikuttavien tekijöiden ja taustaseikkojen variansseja paitsi Z :n (esimerkiksi Dahler-Larsen ja Krogstrup 2003). Väliin tulevat muuttujat, eli moderaattorit mo , ovat tekijöitä joiden suhteen syytekijän vaikutus seuraukseen on erilainen, eli ne vaikuttavat muuttujien välisiin yhteyksiin. Moderaattori määrittää keneen ja miten syy vaikuttaa. Sekoittavat tekijät c ovat yhteydessä sekä syy- että seurausmuuttujaan. Instrumenttimuuttujat i taas vaikuttavat vain käsittelyyn, jolloin ne ovat yhteydessä vasteeseen (seuraus) vain välillisesti käsittelyn kautta. Seuraavassa on esitetty kaavio edellä esitellyistä erilaisista kausaalisuhteeseen vaikuttavista tekijöistä.



Tässä tutkielmassa keskitytään tilanteeseen, jossa sekä käsittely että vaste ovat dikotomisia. Kausaalisuutta pyritään mallintamaan muodostamalla tutkittavista yksilöistä taustamuuttujiltaan mahdollisimman hyvin toisiaan vastaavia pareja huomioimalla sekoittavat tekijät. Käsittelyvaikutusta mitataan vertailemalla vastinparien vasteita kaltaistetussa tilanteessa. Seuraavaksi esitellään satunnaistettua ja havaittua tutkimusasetelmaa tästä näkökulmasta.

3.2 Satunnaistettu koe: ideaalitalanne

Tutkimuksen ideaalityyppinä voidaan pitää Fisherin (1935) esittämää satunnaistettua koetta. Kun yksilöt valitaan satunnaisesti koe- ja kontrolliryhmään, on yksilöillä sama todennäköisyys valikoitua käsittelyyn (Rosenbaum 1995: 17–21), ja todennäköisyys tulee satunnaistettuun asetelmaan mukaan ainoastaan käsittelyn tunnetun satunnaistamismekanismin kautta, ja jäljelle jäävä satunnaisvaihtelu ei vaikuta enää systemaattisesti vasteen arvioihin (Ho ym. 2007: 6; Seidenfeld 1979: 72-76). Kun valikoitumisen mekanismi tunnetaan, on myös satunnaismuuttujan jakauma tunnettu. Koetilanteessa pyritään siihen, että käsittelymuuttujan jakauma tunnetaan, jotta sen vaikutusta voidaan arvioida.

Satunnaistamisen avulla kontrolloidaan taustamuuttujien vaikutusta vasteeseen, sillä riittävän suurissa otoksissa havaitut ja havaitsemattomat muuttujat ovat satunnaistamisen jälkeen likimain samoin jakautuneita käsittelyryhmissä, eli tasapainossa. Tietysti aina on mahdollisuus, että huomiotta jää sattumalta muuttujia, joiden suhteen satunnaistaminen ei tasoita jakaumia käsittelyryhmissä, mutta satunnaistamisen olettavat vaikutuksen tilastolliset mittarit olettavat että satunnaistaminen tasapainottaa ne riittävästi.

Satunnaistamisen toteutumisen ainoa ehto on, että ennen käsittelyyn valikoitumista kaikilla yksilöillä on oltava nollasta poikkeava todennäköisyys valikoitua koe- tai kontrolliryhmään. Yleistettävyyden vuoksi voidaan olettaa, että yksilöt on jaettu ositteisiin s , kun $s = 1, \dots, S$. Ositus tehdään ennen jakoa käsittelyryhmiin. (Osittamista käsitellään myöhemmin myös kaltaistuksen yhteydessä, jolloin verrokkiparit ovat ositteita.) On siis oltava $0 < P(Z_{si}=1) < 1$, kun $Z_{si}=1$ kuvaa ositteen s yksilön i valikoituvan koeryhmään.

Yksinkertaisimmillaan valikoituminen tapahtuu riippumattomasti ja yhtä suurin todennäköisyyksin kaikille yksilöille, jolloin käsittelyn ollessa binäärinen

$$P(Z_{si}=1)=1/2 \text{ jokaiselle } s, i.$$

Nyt N on havaintojen kokonaislukumäärä, eli

$$N = \sum_{s=1}^S n_s .$$

Mahdollisia käsittelyvaihtoehtoja on 2^N kappaletta, ja jokainen erilainen joukko käsittelyjä on yhtä todennäköinen:

$$P(Z_{si} = z_{si}, i = 1, \dots, n_s, s = 1, \dots, S) = \frac{1}{2^N},$$

missä z_{si} :t ovat mielivaltaisia jonoja nollia ja ykkösiä. Tällöin todennäköisyys että kaikki yksilöt valikoituvat samaan käsittelyryhmään on nolasta eroava. Tämä todennäköisyys tosin lähestyy nolaa, kun havaintojen määrä kasvaa. Toisaalta, kun ositteiden määrä S on suuri verrattuna yksilöiden määrään N , on mahdollista että jossain ositteessa s kaikki yksilöt kuuluvat samaan käsittelyryhmään. Tuleekin varmistaa että jokaiseen osajoukkoon tulee varmasti sekä käsiteltyjä että käsittelemättömiä. Jos ositteiden koot n_s ovat vakioisia, $n_s = n$ jokaiselle s , niin käsiteltyjen määrä m_s voidaan valita vakioksi m . Tätä kiinteäositteista koeasetelmaa kutsutaan myös nimellä tasaisesti satunnaistettu koe (*uniform randomized experiment*). Kun $n = 2$ ja $m = 1$, koe on pareittainen satunnaistettu koe (*paired randomized experiment*). (Rosenbaum 1995: 18–19)

Satunnaistamisen ideaali tarkoittaa pareittaisen kokeen kohdalla tilannetta, jolloin jokaisesta parista toinen arvotaan koeryhmään. Tämän jälkeen käsittelyvaikutusta tutkitaan koe- ja kontrolliryhmissä, jotka satunnaistamisen luoman tasapainon johdosta olisivat vertailukelpoisia. Esimerkiksi selvittäessä aktivointisuunnitelmien laadinnan vaikutusta työllistymiseen vastinparit muodostettaisiin etukäteen, ja jokaisesta vastinparista toinen arvottaisiin aktivoitavaksi. Suunnitelmien laadinnan jälkeen seurattaisiin parien työllistymistä. Käsittelyryhmät olisivat vertailukelpoisia, sillä ne olisivat satunnaistamisen johdosta taustamuuttujiltaan toisiaan vastaavia. Työttömät eivät kuitenkaan todellisuudessa valikoituneet aktivoitaviksi satunnaisesti.

3.3 Käsittelyvaikutus ja testaus satunnaistetussa koetilanteessa

3.3.1 Käsittelyvaikutus

Jos käsittely vaikuttaa vasteeseen, ovat yksilöiden vasteiden arvot erilaisia käsittelyn eri tasoilla (Rosenbaum 1995: 32). Syy-seuraussuhteen tutkiminen tarkoittaa Rubinin ja Rosenbaumin lähestymistavassa saman yksilön vasteiden vertailua koe- ja kontrollitilanteessa. Todellisuudessa tämä ei ole yleensä mahdollista, vaan yksilö valikoituu joko koe- tai kontrolliryhmään.

Potentiaaliset tulosvaihtoehdot kuvaavat yksilön mahdollisia vasteen arvoja eri käsittelyryhmissä ennen käsittelyyn valikoitumista. Vaikutukset ovat potentiaalisten vasteiden funktioita. Merkitään yksilön i potentiaalisia tulosvaihtoehtoja joko $(Y_i | T_i = 0)$ tai $(Y_i | T_i = 1)$ kontrolli- ja koetilanteessa (Ho ym. 2007: 4). Yksilön i havaittu vaste y_i on realisaatio vastaavasta satunnaismuuttujasta Y_i . Potentiaalisten vasteiden erotuksen odotusarvoa kutsutaan käsittelyvaikutukseksi (Ho ym. 2007: 4). Merkitään t :lla satunnaisen kausaalisen vaikutuksen odotusarvoa

$$t_i = E[(Y_i | T_i = 1) - (Y_i | T_i = 0)].$$

Koska potentiaalisista vasteista toteutuu vain toinen, on käsittelyvaikutus todellisuudessa tuntematon. (Rosenbaum 1995). Työllistymisesimerkissä potentiaalisten vasteiden erotuksella mitattaisiin mahdollisen aktivointisuunnitelman laadinnan vaikutusta työllistymiseen ennen aktivointiin valikoitumista. Todellinen käsittelyvaikutus on tuntematon, joten käsittelyvaikutuksen estimoinnissa on huomioitava aktivointiin valikoitumisen mekanismi.

Yleisesti käsittelyvaikutusta t voidaan mitata käsittelyryhmien havaittujen vasteiden keskiarvojen erotuksilla. Käsittelyvaikutukselle oletetaan tässä additiivinen malli, jolloin t saadaan koe- ja kontrolliryhmän vasteiden erotuksen kautta (Rosenbaum 1995: 34–35). Tärkeä oletus satunnaistetuissa kokeissa on, että ne johtavat keskimääräisen käsittelyvaikutuksen harhattomiin estimaatteihin. Yksinkertaisimmillaan kokeessa on vain yksi osite, ja havainnot oletetaan riippumattomiksi. Tässä tasaisesti satunnaistetussa kokeessa ($m_s = m$) koeyksilöiden määrä on m ja

kontrollien vastaavasti $N - m$. Tällöin $E(Z_i) = m/N$, $E(1 - Z_i) = 1 - m/N$, ja keskimääräinen käsittelyvaikutus on keskiarvo havaittujen vasteiden erotuksista

$$\frac{1}{N} \sum_i (y_{Ti} - y_{Ci}),$$

jossa y_{Ti} on koeryhmän i :nnen yksilön vaste, ja y_{Ci} kontrolliryhmän i :nnen yksilön vaste. Käsittelyvaikutuksen $y_{Ti} - y_{Ci}$ ei edellä oleteta olevan kiinteä. Kun oletetaan käsittelyvaikutuksen olevan additiivista, on vaikutus positiivista jos $t > 0$ ja $y_{Ti} \geq y_{Ci}$.

Tämä käsittelyvaikutuksen estimaattori estimoi keskimääräistä vaikutusta harhattomasti ainoastaan kun asetelma on satunnaistettu. Se antaa harhaisen tuloksen jos koeyskilöitä on eri määrä mahdollisissa eri osajoukoissa (Rosenbaum 1995: 37.). Mahdollinen ositus tulee ottaa huomioon siten, että lasketaan erotus koe- ja kontrolliryhmän keskimääräisen vasteen välillä, ja painotetaan tämä erotus yksilöiden lukumäärällä osajoukossa s . Osituksen huomioiva keskimääräisen käsittelyvaikutuksen estimaattori on harhaton (Rubin 1977), sillä voidaan osoittaa, että sen odotusarvo antaa saman estimaattorin kuin edellä esitetty, sillä tällöin satunnaistetussa kokeessa Z_{si} :n odotusarvo on m_s/n_s . (Rosenbaum 1995: 38)

3.3.2 Summatestit ja McNemarin testisuure

Pareittaisessa satunnaistetussa kokeessa, kun kaltaistettu aineisto muodostuu S :stä parista ja vaste on binäärinen, käytetään McNemarin testiä. McNemarin testisuure on erikoistapaus Mantel-Haenzel testisuureesta jota käytetään usein kaltaistetuissa epidemiologisissa asetelmissa, kun jokaiselle koeryhmän yksilölle on useampia verrokkeja. McNemarin testissä ositteet ovat koe-verrokki - pareja. Vastinparien frekvenssitaulukko on esitetty seuraavassa. (Rosenbaum 1995: 21-32).

		koeryhmä		total
		0	1	
kontrolliryhmä	0	f_{00}	f_{01}	$f_{0.}$
	1	f_{10}	f_{11}	$f_{1.}$
		$f_{.0}$	$f_{.1}$	S

Kontrolliryhmästä positiivisen vasteen arvon saaneita on $f_{1.}$ yksilöä, ja koeryhmästä vastaavasti $f_{.1}$. Frekvenssit f_{01} ja f_{10} kuvaavat pareja, joissa vain toisella parissa on positiivinen vasteen arvo. Kun y_T on koeyksilön vaste, ja y_C vastaavasti kontrolliyksilön vaste, merkitään koe- ja kontrolliyksilöiden positiivisten vasteiden arvojen todennäköisyyksiä

$$q_T = P(y_T = 1) \text{ ja } q_C = P(y_C = 1).$$

Edellinen taulukko voidaan nyt esittää todennäköisyyksien avulla:

		koeryhmä		total
		0	1	
kontrolliryhmä	0	p_{00}	p_{01}	$1 - q_C$
	1	p_{10}	p_{11}	q_C
		$1 - q_T$	q_T	1

Todennäköisyys että parin koeyksilön vaste ei ole positiivinen on $(1 - q_T)$ ja vastaavasti kontrolleille $(1 - q_C)$, jolloin todennäköisyyksien summa on 1. Nyt p_{01} on todennäköisyys että parin kontrolliyksilöllä vaste saa positiivisen arvon, ja koeyksilöllä ei. Samoin p_{10} on todennäköisyys, että parin koeyksilöllä vaste on positiivinen ja kontrolliyksilöllä taas ei.

Vastinparien frekvenssitaulukossa marginaalit voivat nyt olla mitä vain. Nyt f_{01} noudattaa ehdollista binomijakaumaa ehdolla $f_{01} + f_{10}$. Voidaan osoittaa, että (mm. Tuominen 1993)

$$f_{01} | f_{01} + f_{10} \sim \text{Bin} \left(\frac{p_{01}}{p_{01} + p_{10}}, f_{01} + f_{10} \right).$$

Nollahypoteesitilanteessa, jolloin käsittely ei vaikuta vasteiden arvoihin, eivät ne frekvenssit, joissa parin yksilöillä on eriävät vasteen arvot, eroa toisistaan. Tällöin todennäköisyydet $p_{01} = p_{10}$, eli $p_{01} / (p_{01} + p_{10}) = 1/2$ ja valikoitumisen todennäköisyys yksilöllä i on $p_i = 1/2$. Näin ollen

$$f_{01} | f_{01} + f_{10} \sim \text{Bin} \left(\frac{1}{2}, f_{01} + f_{10} \right).$$

Aktivointiesimerkissä koeryhmä muodostuu aktivoituista, ja vasteen positiiviset arvot vastaavat työllistymistä. Vastinparien frekvenssitaulukoilla merkitsevyyttä voidaan testata McNemarin testillä, jolloin merkitsevyyttä testataan ehdollista binomijakaumaa vasten.

4. Harha ja havainnoiva tutkimus

4.1 Havainnoivan tutkimuksen ero satunnaistettuun koetilanteeseen

Satunnaistetun kokeen ideaalitalanteessa käsittelyryhmät ovat vertailukelpoisia. Havainnoivassa tutkimuksessa valikoituminen koeryhmään ei tapahdu satunnaisesti. Valikoituminen voidaan huomioida esimerkiksi osituksen tai kaltaistuksen avulla ennen tilastollista analyysiä. Oletetaan nyt että x_i on sekoittava taustatekijä yksilölle i , ja että yksilön valikoitumista koe- tai kontrolliryhmään kuvaa $Z_i = 1$ käsittelyn saaneille ja $Z_i = 0$ kontrolleille. Kun yksilöt ovat valikoituneet käsittelyyn satunnaisesti ja toisistaan riippumatta, käsittelytodennäköisyyksien malli on tulo yksilöiden käsittelytodennäköisyyksistä

$$P(Z_1 = z_1, \dots, Z_N = z_N) = \prod_{i=1}^N p_i^{z_i} \{1 - p_i\}^{1-z_i},$$

missä $Z_i = 0$ tai 1 ja käsittelytodennäköisyydet $p_i = P(Z_i = 1)$ ovat tuntemattomia. Havainnoivassa tutkimuksessa käsittelytodennäköisyys p voi olla erisuuri eri yksilöillä, jolloin yksilöillä ei ole sama todennäköisyys valikoitua käsittelyyn, ja mallissa voi olla valikoitumisharhaa. (Rosenbaum 1995: 61-69)

Esimerkiksi tutkittaessa aktivoinnin vaikutusta työllistymiseen, aktivointiin valikoituminen riippui niin yksilön piirteistä, työttömyyden pituudesta kuin sosiaalityöntekijän priorisoineista sen suhteen kuka kipeimmin aktivointia tarvitsi: aktivointitodennäköisyydet olivat erisuuruisia. Valikoitumisharhan vuoksi käsittelyryhmät eivät olleet vertailukelpoisia kaikkien kovariaattien suhteen, joten oli mahdollista, että koe- ja kontrolliryhmät erosivat toisistaan jo ennen käsittelyyn valikoitumista. Valikoitumisen mekanismia ei tällöin tunneta. Tässä havainnoivassa tutkimusasetelmassa lähtökohtana onkin joukko aktivoituja, joita tulisi verrata taustatekijöiltään vastaavaan ei-aktivoitujen työttömien ryhmään.. Havainnoivissa tutkimusasetelmissä onkin aina huomioitava mahdollisen valikoitumisharhan vaikutus tuloksiin.

Jos valikoituminen huomioidaan asetelmassa ja näin jäljitellään satunnaistettua tilannetta, voidaan käsittelyn vaikutusta tutkia suoraan satunnaistamisen olettavilla menetelmillä. Kaltaistettujen parien

vasteiden arvojen avulla estimoidaan tällöin tuntematonta potentiaalisten vasteiden erotusta. Jos koe ei ole satunnaistettu, tarvitaan oletus jakaumasta $P(Z = z)$ (Rosenbaum 1995: 22). Jotta käsittelyvaikutus voitaisiin estimoida harhattomasti, on löydettävä tapoja hallita valikoitumisharhaa.

4.2 Harha

Havainnoivissa tutkimuksissa harha on tilastollista epätasapainoa tutkittavien ryhmien välillä. Havainnoivassa tutkimusasetelmassa voi olla havaittavaa harhaa (*overt bias*) tai havaitsematonta eli piilevää harhaa (*hidden bias*). Jos koe- ja kontrolliryhmä eroavat toisistaan systemaattisesti ennen käsittelyä, on mukana valikoitumisharhaa, johon vaikuttavat niin havaittava kuin havaitsematonkin harha. Havaittava harha voidaan estimoida mitatuista muuttujista x . Kausaalisuhteeseen saattavat vaikuttaa myös havaitsemattomat tekijät joista ei aineistossa ole tietoa. Valikoitumisharhan minimoimiseksi on niiden olemassaolo tunnistettava.

Havaittavan harhan suuruutta voidaan arvioida vertailemalla kovariaattien x jakaumia käsittelymuuttujan eri ryhmissä. Vasteeseen vaikuttavien tarpeeksi suurien erojen havaitsemiseen tarvitaan tilastollisia testejä. Yksi tapa löytää epätasapainoiset kovariaatit on vertailla koe- ja kontrolliryhmiä eri kovariaattien suhteen t-testien avulla. Tämän jälkeen kovariaateilla ehdollistamisessa käytetään vain niitä kovariaatteja joilla löytyy merkitsevä ero ryhmien välillä. (Rosenbaum 1995: 61-64.) Tähän liittyy Rosenbaumin (1995: 64) mukaan kuitenkin kolme ongelmaa. Ensinnäkään näin ei saada tietoa kovariaatin ja vasteen suhteesta, eikä toisaalta voida perustellusti väittää, että tilastollisen merkitsevyyden puuttuminen tarkoittaisi kovariaattien riittävää tasapainoa. Kolmanneksi näin verrataan käsittelyryhmiä keskenään vain yhden kovariaatin suhteen kerrallaan, kun tulisi voida huomioida useita samanaikaisesti. Tämä voi kuitenkin olla lähtökohtana tieteellisesti perusteltujen kovariaattivalintojen ohella, kunhan ehdollistamisen jälkeen tarkastetaan kovariaattien tasapaino uudelleen. Kun kovariaattien jakaumat ovat samankaltaisia käsittelyryhmissä, ovat ryhmät vertailukelpoisia keskenään niiden osalta.

Jos mukana on havaitsematonta eli piilevää harhaa, ei voida sanoa, että mahdollinen havaittu ero käsittelyryhmien vasteiden välillä johtuisi käsittelyn vaikutuksesta. Tällöin vaikka yksilöiden havaittujen kovariaattien arvot olisivat samat, eivät niiden käsittelytodennäköisyydet välttämättä ole, eli

$$x_i = x_j, \text{ mutta mahdollisesti } p_i \neq p_j.$$

Kaltaistuksen ja osituksen tavoitteena ei olekaan vain luoda vertailuryhmiä jotka ovat homogeenisia kovariaattien suhteen. Jos havaitsematonta harhaa ei ole, riittää ehdollistaa havaittujen kovariaattien suhteen. Tämän jälkeen analysoinnissa ja käsittelyvaikutuksen tilastollisessa testauksessa on mahdollista käyttää menetelmiä jotka perustuvat satunnaistamiseen. Toisaalta, riippumatta havaitsemattoman harhan olemassaolosta, on mahdollista luoda vertailukelpoiset ryhmät esimerkiksi käsittelyn suhteen vaikka yksilöillä olisi eri kovariaattien arvot. Tällöin pyritään kovariaattien tasapainoon, jolloin kovariaateilla on vertailuryhmissä samankaltaiset jakaumat. (Rosenbaum 1995: 88, 200)

4.3 Havaitun valikoitumisharhan kontrollointi

Havaitun harhan vaikutus voidaan huomioida osituksen tai kaltaistuksen avulla. Tällöin käytetään aineistossa olevaa tietoa, jonka avulla on mahdollista vähentää valikoitumisharhaa. Saadut vaikuttavuusestimaatit kuvaavat käsittelyn vaikutusta kun havaittujen taustamuuttujien vaihtelu on vakioitu. Myös tilastollisella mallinnuksella voidaan tutkia eri ositteiden sisällä käsittelyvaikutusta, jolloin osituskovariaattia käytetään ehdollistavana muuttujana.

4.3.1 Ositus

Ositus on yksinkertaisin ja karkein tapa huomioida havaittu harha. Tällöin otos jaetaan osiin sekoittavan kovariaatin x suhteen. Kovariaatin erillisistä luokista s poimitaan käsiteltyjä ja käsittelemättömiä yhteensä n_s kappaletta. Tällöin osituskovariaatin on oltava diskreetti (tai mataladimensioinen). (Rosenbaum 1995: 66-67)

Kun ositetaan täsmälleen, on yksilöillä oltava sama kovariaatin x arvo jotta ne kuuluisivat samaan ositteeseen, jolloin $x_{si} = x_{sj}$. Jos aineistossa ei ole piilevää harhaa, on joka ositteen sisällä olevilla yksilöillä osituksen jälkeen sama käsittelytodennäköisyys p . Ositus voidaan tehdä myös optimaalisesti, jolloin koe ja kontrolliyksilöiden väliset etäisyydet osituskovariaattien x suhteen ovat ositteissa mahdollisimman pieniä. Näin ollen myös kaltaistus on eräänlaista ositusta. (Rosenbaum 1995: 205-206)

4.3.2 Kaltaistus

Kaltaistuksella (*matching*) pyritään havainnoivassa tutkimusasetelmassa satunnaistettua koeasetelmaa vastaavaan tilanteeseen. Kaltaistuksessa koeyksilöille etsitään havaittujen kovariaattien, tai niiden funktion, suhteen vastaava verrokki. Tällöin sekoittavan tekijän jakauma on mahdollisimman samankaltainen eri käsittelyryhmissä. Jos kovariaatin suhteen kaltaistetaan, ei kyseessä olevasta kovariaatista voida tehdä päätelmiä analyysivaiheessa, sillä kaltaistuksessa kovariaatit tasapainotetaan niin, ettei niiden suhteen ole eroa käsittelyryhmittäin. Näin luodussa tilanteessa jäljitellään yksilön potentiaalisten vasteiden eroa samankaltaisten yksilöiden realisoituneiden vasteiden eron avulla.

Takautuvissa (*case-control*) asetelmissa tapa huomioida yksilöiden väliset erot on muodostaa pareja, joissa kaltaistetut yksilöt ovat eri käsittelyryhmistä. Tällöin koe- ja kontrolliryhmissä on yhtä monta yksilöä, ja estimointi perustuu hypergeometriseen jakaumaan, taulukon reuna- eli ehdollisten jakaumien mukaisesti (Clayton ja Hills 1993: 331–332). Jos kaltaistetaan suoraan havaittujen muuttujien perusteella, etsitään koeyksilölle verrokki jolla on samat muuttujien arvot. Koska täydellisesti samankaltaisia pareja tai ryhmiä voi olla mahdotonta löytää varsinkin jos kaltaistetaan useiden muuttujien suhteen, on hyödyllistä joustaa vaatimuksissa ja etsiä verrokki jonka muuttujien arvot ovat heti viereisissä luokissa. On tasapainoiltava kahden ääripään välillä, toisaalta että löytyisi mahdollisimman paljon täydellisesti yhteensopivia (eli kaltaistusmuuttujien suhteen samanlaisia) yksilöitä, jolloin saattaa jäädä paljon yksilöitä ilman kaltaista verrokkia, sekä toisaalta että löytyisi mahdollisimman paljon pareja, jolloin on hyväksyttävä epätarkka kaltaistus (ks. Aho ja Kunttu 2001).

Mitä useampi verrokki, sitä paremmin saadaan tutkittavia parametreja estimoitua. Jos $n_s \geq 2$ ja $m_s = 1$ kaikille $s = 1, \dots, S$ niin jokaiselle käsittelyn saaneelle yksilölle on asetelmassa vähintään 2 kontrollia (Rosenbaum 1995: 18). Verrokkeja kannattaa olla useampi kuin yksi, enintään kuitenkin 5. Jos oletetaan, että verrokkeja on v kertaa niin paljon kuin tapauksia, niin tapaus-verrokki tutkimuksen tarkkuus verrattuna kohorttitutkimukseen on karkeasti arvioituna $\sqrt{1 + (1/v)}$, jolloin mainittavaa etua ei kerry enää viidennen kontrollin jälkeen. (Clayton ja Hills 1993: 331–332)

On myös muita tapoja vertailla vasteita koe- ja kontrolliryhmissä kaltaistuksen avulla. Harvemmin kuitenkaan poimitaan useita koeyksilöitä yksilöitä per kontrolli. Toisaalta on mahdollista erotella koeryhmälle erilaisia verrokkiryhmiä, jolloin kontrolliryhmiä on useita erilaisia. Voidaan myös valita vertailuun erilaisia viiteryhmiä, joissa jokaisessa on tapauksia ja verrokkeja. (Rosenbaum 1995: 154–155)

Optimaalisinta on täysi kaltaistus. Optimaalisuus määritellään tässä kaltaistettujen ryhmien väliseksi minimoiduksi etäisyydeksi. Täysin kaltaistetussa otoksessa jokainen kaltaistettu joukko sisältää joko yhden koeyksilön ja yhden tai useamman verrokin, tai yhden verrokin ja yhden tai useamman koeyksilön. Parikaltaistus ei ole optimaalista, eikä yleisemmin kaltaistus jossa verrokkien määrä on kiinteä. Tämä johtuu siitä, että koe- ja kontrolliryhmä ovat jakautuneet eri tavalla havaittujen kovariaattien x suhteen, jolloin toisilla kovariaattien arvoalueilla on enemmän koeyksilöitä, toisilla enemmän kontrolleja. (Rosenbaum 1995: 201)

4.3.3 Propensiteettipistemäärä

Propensiteettipistemääräksi (*propensity score*) $x(x)$ kutsutaan kovariaattien x funktiota, joka ennustaa todennäköisyyttä kuulua käsittelyryhmään havainnoivassa tutkimuksessa (mm. Rosenbaum ja Rubin 1983). Tällä funktiolla korvataan kaltaistuksessa joukko sekoittavia kovariaatteja (Rubin 1997: 760). Menetelmä on hyödyllinen kun sekoittavia taustamuuttujia on useita, ja osa niistä jatkuvia. Olennainen etu on, että kaltaistus tehdään yhden pistemäärämuuttujan avulla, eikä tukeuduta useisiin muuttujiin. Se on myös moniulotteinen lähestymistapa verrattuna esimerkiksi ositukseen yhden muuttujan suhteen. Sen käyttö vaatii suurta otoskokoa. (Rosenbaum 1995: 69-70)

Propensiteettipistemäärän estimointiin käytetään standardeja todennäköisyysmalleja (Deheija ja Wahba 2002: 161), esimerkiksi logistista regressiota tai diskriminanattianalyysiä (Rubin 1997: 760). Selitettävänä muuttujana on dikotominen käsittelymuuttuja ja selittävinä muuttujina ovat sekoittavat tekijät. Haaste liittyykin käsittelytodennäköisyyden ennustamiseen eikä estimointiin. Kun ennustettu käsittelyyn valikoitumisen todennäköisyys saadaan estimoitua kaikille yksilöille, käytetään sitä selittäväksi muuttujan mallinnettaessa käsittelyn vaikutusta.

Oletetaan, että propensiteettipistemäärä tunnetaan. Propensiteettipistemäärää voidaan käyttää kahdella tapaa. Rubin (1997: 760) suosittelee, että propensiteettipistemäärämuuttuja olisi hyvä luokitella noin viiteen luokkaan. Tämän jälkeen riittäisi kaltaistaa tämän luokitellun todennäköisyysmuuttujan suhteen. Toinen mahdollinen tapa on lähimmän naapurin menetelmä, jossa etsitään propensiteettipistemäärältään lähinnä olevat koe- ja verrokkiyksilöt vertailtaviksi. Propensiteettipistemäärää voidaan käyttää kaltaistuksessa tai osituskovariaattina yksin tai toisten kaltaistusmuuttujien kanssa.

Jos tiedetään, että tuntemattomat käsittelytodennäköisyydet $0 < p_i < 1$ riippuvat ainoastaan tunnetuista kovariaateista, ei kokeessa ole piilevää harhaa. Tällöin

$$p_i = x(x_i),$$

ja on riittävää kaltaistaa tai osittaa propensiteettipistemäärän suhteen. Käsittelymuuttujan yhteistodennäköisyys voidaan esittää propensiteettipistemäärän avulla muodossa

$$P(Z_1 = z_1, \dots, Z_N = z_N) = \prod_{i=1}^N x(x_i)^{z_i} \{1 - x(x_i)\}^{1-z_i}.$$

(Rosenbaum 1995: 65-66.) Tätä kutsutaan satunnaistamiseksi kovariaatin perusteella (Rubin 1977).

Esimerkiksi logistisessa regressiossa ennustettu todennäköisyys

$$x(x) = \frac{e^{x'B}}{1 + e^{x'B}},$$

jossa X on havaittu muuttujavektori ja B on estimoitava parametrivektori. Kun saadaan estimoitu \hat{B} , kaltaistuksessa käytettäväksi propensiteettipistemääräksi muodostuu

$$\hat{X}(x) = \frac{e^{x'\hat{B}}}{1 + e^{x'\hat{B}}}.$$

Kun funktiota $\hat{X}(x)$ käytetään kaltaistuksessa, eroavat sekoittavien kovariaattien jakaumat vertailuryhmissä vain satunnaisvaihtelultaan. Tällöin on mahdollista käyttää satunnaistamiseen perustuvia käsittelyvaikutuksen t estimointimenetelmiä olettaen, että todellinen propensiteettipistemäärä on mahdollista estimoida, ja että piilevää harhaa ei ole. Toisaalta vaikka aineistossa olisikin piilevää harhaa, propensiteettipistemäärällä ositus tai kaltaistus tasapainottaa kovariaatit x , kuten satunnaistetussa koetilanteessa. Piilevien kovariaattien epätasapainottavaa vaikutusta tämä ei kuitenkaan poista. (Rosenbaum 1995: 201-204)

4.4 Sensitiivisyysanalyysi ja piilevän harhan vaikutuksen estimointi

Piilevän harhan vaikutusta tutkimustuloksiin arvioidaan sensitiivisyysanalyysillä. Sensitiivisyysanalyysin avulla voidaan osoittaa miten paljon havaitsemattomat tekijät vaikuttaisivat tuloksiin, mutta se ei kerro onko näitä tekijöitä todella.

Jos esimerkiksi erilaiset kontrolliryhmät eroavat vasteiltaan systemaattisesti, ei syynä voi olla käsittely, vaan juuri havaitsematon harha. Myös sellainen systemaattinen ero koe- ja kontrolliryhmien vasteiden välillä johon käsittely ei vaikuta, on merkki havaitsemattoman harhan vaikutuksesta. Kun tutkitaan havaitsemattomien muuttujien u vaikutusta, tutkitaan näkykö käsittelyn vaikutus siellä missä sen pitäisi, eikä muualla. Tilastollisia välineitä piilevän harhan vaikutuksen arviointiin eivät ole ainoastaan havaitut tekijät vaan myös useat referenssiryhmät tai useat kontrolliryhmät (Rosenbaum 1995: 136-137).

Vaativamman suuruista harhaa voi olla vaikea tai mahdotonta osoittaa aineiston avulla, mutta toisaalta sensitiivisyysanalyysin avulla on mahdollista todeta vähäisen harhan mahdollinen

vaikutus. Kaikkien kokeiden tulokset muuttuvat jos harhaa on tarpeeksi paljon, mutta suuri harha on myös helpompi osoittaa.

Jos tutkimuksessa on havaitsematonta harhaa, voivat käsittelytodennäköisyydet vaihdella havaitsemattoman kovariaatin suhteen. Tällöin kahdella yksilöllä joilla on samat kovariaattien x arvot, ei ole sama todennäköisyys joutua käsittelyryhmään, eli vaikka $x_i = x_j$, niin $p_i \neq p_j$. Kaltaistettaessa oletus oli, että havaituilla kaltaistavilla tekijöillä pystytään kuvaamaan käsittelyyn valikoituminen. Jos havaitsematonta harhaa on mukana, eivät oletukset toteudu. (Rosenbaum 1995: 88)

4.4.1 Kokeen sensitiivisyys

Sensitiivisyysanalyysillä pyritään kuvaamaan kuinka suuria käsittelytodennäköisyyksien p_i erojen on oltava havaittujen kovariaattien eri luokissa, jotta ne vaikuttavat tutkimuksen tuloksiin. Toisin sanoen tutkitaan, kuinka suuria tulisi piilevien kovariaattien vedonlyöntisuhteiden olla, jotta havaittu tulos ei olisi tilastollisesti merkitsevä. Kahdelle yksilölle, i ja j , käsittelyyn valikoitumisen vedonlyöntisuhteet ovat

$$p_i/(1-p_i) \text{ ja } p_j/(1-p_j).$$

Jos oletetaan, että molemmin päin lasketut ristitulosuhteet yksilöille joille $x_i = x_j$ olisivat enintään $\Gamma \geq 1$, niin

$$\frac{1}{\Gamma} \leq \frac{p_i/(1-p_i)}{p_j/(1-p_j)} = \frac{p_i(1-p_j)}{p_j(1-p_i)} \leq \Gamma,$$

kaikille i, j (Rosenbaum 1995: 88). Jos nollahypoteesitilanne $\Gamma=1$ on voimassa, niin $p_i = p_j$ kun $x_i = x_j$, jolloin mukana ei ole havaitsematonta harhaa. Käsittelyyn valikoituminen on siis yhtä todennäköistä yksilöille huomioitaessa x . Jos käsittelyvaikutus on ollut merkitsevä,

sensitiivisyysanalyysissä tutkitaan havaitsemattomien tekijöiden aiheuttaman vedonlyöntisuhteen suuruuden vaikutusta tulkintoihin useiden eri Γ :n arvojen avulla.

Γ :n avulla voidaan tutkia mahdollisen piilevän harhan vaikutusta, sillä vertailussa havaitut kovariaatit häviävät. Oletetaan, että yksilöön i liittyy sekä havaittu kovariaatti x_i sekä piilevä kovariaatti jota merkitään u_i . Kun havaitsemattomat tekijät u otetaan huomioon mallissa, saadaan vedonlyöntisuhteen logaritmilille (logit) funktio

$$\log\left(\frac{p_i}{1-p_i}\right) = k(x_i) + gu_i,$$

jossa k on tuntematon funktio ja g on tuntematon parametri. Oletetaan seuraavassa, että $0 \leq u_i \leq 1$. Jos lisäksi oletetaan että $x_i = x_j$ yksilöille i ja j , ja siten myös $k(x_i) = k(x_j)$, jolloin nämä yksilöt saattaisivat valikoitua kaltaisiksi. Nyt ristitulosuhde yksilöille i ja j valikoitua koeryhmään on

$$\frac{p_i(1-p_j)}{p_j(1-p_i)} = \exp\{g(u_j - u_i)\},$$

joka ei enää riipu x :stä. Näin ollen yksilöt joilla on sama havaitun kovariaatin arvo x , eroavat käsittelyyn valikoitumisen ristitulosuhteeltaan havaitsemattomien kovariaattien erotuksen funktion kautta.

Kun oletetaan että $e^g = \Gamma \geq 1$ ja $0 \leq u_i \leq 1$, on $-1 \leq u_j - u_i \leq 1$ ja siten $\max_{i,j} e^{g(u_j - u_i)} = e^g$, ja vastaavasti $\min_{i,j} e^{g(u_j - u_i)} = e^{-g}$. Yksilöiden i ja j ristitulosuhde valikoitua koeryhmään rajoitetaan näin välille $[e^{-g}, e^g]$. (Rosenbaum 1995: 89–90)

Koe on sensitiivinen, tai herkkä, jos tulkinnat muuttuvat olennaisesti kun Γ on lähellä 1:stä. Tällöin jo pienet poikkeamat nollahypoteesitilanteesta johtaisivat erilaisiin tulkintoihin. Jos tulkinnat muuttuisivat vasta äärimmäisillä Γ :n arvoilla, koe ei ole herkkä. Jos tulkinnat muuttuisivat esimerkiksi vasta kun $\Gamma = 6$, saattaisi kaksi havaittujen tekijöiden suhteen samankaltaista yksilöä

(eli yksilöä joille $x_i = x_j$) erota käsittelyyn joutumisen ristitulosuhteeltaan kuusinkertaisesti. Tällaisen tekijän olemassaolo on jo epätodennäköistä. Γ voidaan näin ollen määritellä mitaksi, joka kertoo kuinka kaukana tilanne on kokeesta, jossa piilevää harhaa ei ole. (Rosenbaum 1995: 95)

Jos aineisto ositetaan muuttujan x suhteen S :ään ositteeseen, jolloin ositteessa s on n_s havaintoa, joista m_s saa käsittelyyn, ei valikoitumisen mekanismin jakauma $Z = (Z_{11}, \dots, Z_{S, n_s})$ annetulla m ole enää kiinteä kuten tapauksessa jolloin piilevää harhaa ei ole. Sen sijaan sen jakaumaksi voidaan johtaa

$$P(Z = z | m) = \prod_{s=1}^S \frac{\exp(\mathbf{g}_{z_s}^T \mathbf{u}_s)}{\sum_{z_s \in \Omega_s} \exp(\mathbf{g}_{z_s}^T \mathbf{u}_s)},$$

jossa $\mathbf{z}_s = [z_{s1}, \dots, z_{sn_s}]^T$, $\mathbf{u}_s = [u_{s1}, \dots, u_{sn_s}]^T$, ja Ω_s on joukko joka sisältää $\binom{n_s}{m_s}$ erilaista jonoa, joissa on m_s ykköstä ja $n_s - m_s$ nollaa. (Rosenbaum 1995: 91)

Tällöin tietyllä m , käsittelyyn valikoitumisen Z jakauma ei enää riipu tuntemattomasta funktiosta $k(x)$, mutta riippuu edelleen tuntemattomista kovariaateista u . Ainoastaan jos $\mathbf{g} = 0$, jolloin havaitsematon tekijä u ei ole olennainen käsittelyyn valikoitumisen suhteen, tai jos $u_{si} = u_{sj}$ kaikille $1 < i < j < n_s$ ja $s = 1, \dots, S$, jolloin kaikilla yksilöillä ositteessa s havaitsematon u saa saman arvon, analyysi ei riipu u :sta. Molemmissa tapauksissa jakauma johtaa tulokseen $1/K$, joka on todennäköisyys kaikille K käsittelyyn valikoitumisille Ω :ssa. (Rosenbaum 1995: 91)

4.4.2 Merkitsevyytasojen sensitiivisyys

Tässä osiossa käsitellään Rosenbaumin (1995: 92-93) sensitiivisyysanalyysiä kaltaistetuille pareille. Olkoon kaltaistettuja pareja S kappaletta, jolloin jokaisessa ositteessa $s = 1, \dots, S$ on $n_s = 2$ yksilöä. Parista toinen kuuluu koeryhmään, eli $1 = m_s = Z_{s1} + Z_{s2}$, jolloin $Z_{s2} = 1 - Z_{s1}$. Malli kuvaa S :ää riippumatonta binääristä koetta todennäköisyyksillä

$$P(Z = z | m) = \prod_{s=1}^S \left[\frac{\exp(gu_{s1})}{\exp(gu_{s1}) + \exp(gu_{s2})} \right]^{z_{s1}} \left[\frac{\exp(gu_{s2})}{\exp(gu_{s1}) + \exp(gu_{s2})} \right]^{1-z_{s1}}.$$

Satunnaistetussa koetilanteessa, kun $g = 0$ ja jokaisella yksilöllä on sama käsittelyyn valikoitumisen todennäköisyys, palautuu malli tasajakaumaan. Jos $g > 0$, on parin s ensimmäisellä yksilöllä suurempi todennäköisyys valikoitua käsittelyyn kuin parin toisella yksilöllä jos $u_{s1} > u_{s2}$.

Testisuure kaltaistetuille pareille on

$$T = t(Z, y) = \sum_{s=1}^S \sum_{i=1}^2 c_{si} Z_{si},$$

jossa binäärinen c_{si} on vasteen arvo ja Z_{si} kertoo kumpi yksilö parissa on käsitelty. Testisuureen jakauma on tuntematon, sillä käsittelyyn valikoitumisen Z jakaumassa (g, u) on tuntematon. Sensitiivisyysanalyysissä tarkastellaan tilannetta, jossa $g > 0$, kun nollahypoteesitilanteessa $k = 0$ ja $g = 0$. Jokaiselle mahdolliselle (g, u) :lle testisuure $t(Z, y)$ on summa S :stä riippumattomasta satunnaismuuttujasta, jossa s :s muuttuja saa arvon 1 todennäköisyydellä

$$p_s = \frac{c_{s1} \exp(gu_{s1}) + c_{s2} \exp(gu_{s2})}{\exp(gu_{s1}) + \exp(gu_{s2})},$$

ja arvon 0 todennäköisyydellä $(1 - p_s)$. Kaltaistettu pari on yhtenevä jos $c_{s1} = c_{s2}$, ja epäyhtenevä jos $c_{s1} \neq c_{s2}$. Jos $c_{s1} = c_{s2} = 1$ niin $p_s = 1$, ja jos $c_{s1} = c_{s2} = 0$ niin $p_s = 0$. Näin ollen yhtenevät parit

vaikuttavat testisuureeseen kiinteillä arvoilla kaikilla mahdollisilla (g, u) , ja ainoastaan epäyhtenevät parit vaikuttavat testisuureen merkitsevyyteen.

Vaikka testisuureen nollahypoteesijakauma on tuntematon, niin jokaiselle kiinteälle g nollahypoteesijakaumaa rajoittavat kaksi tunnettua jakaumaa. Kun $\Gamma = e^g$, asetetaan

$$p_s^+ = \begin{cases} 0 & \text{jos } c_{s1} = c_{s2} = 0 \\ 1 & \text{jos } c_{s1} = c_{s2} = 1 \\ \frac{\Gamma}{1+\Gamma} & \text{jos } c_{s1} \neq c_{s2} \end{cases}$$

$$p_s^- = \begin{cases} 0 & \text{jos } c_{s1} = c_{s2} = 0 \\ 1 & \text{jos } c_{s1} = c_{s2} = 1 \\ \frac{1}{1+\Gamma} & \text{jos } c_{s1} \neq c_{s2} \end{cases}$$

joissa $p_s^- < p_s < p_s^+$, ja c_{si} määräytyvät aineiston perusteella. Olkoon T^+ S :n riippumattoman satunnaismuuttujan summa siten että s :s muuttuja saa arvon d_s todennäköisyydellä p_s^+ , ja arvon 0 todennäköisyydellä $1 - p_s^+$. T^- määritellään vastaavasti todennäköisyydellä p_s^- . Kaikille $u \in U$ testisuureen tuntematon jakauma voidaan rajoittaa näiden kahden testisuureen, T^- ja T^+ , jakaumien avulla. Jos käsittelyllä ei ole vaikutusta, niin jokaiselle kiinteälle g

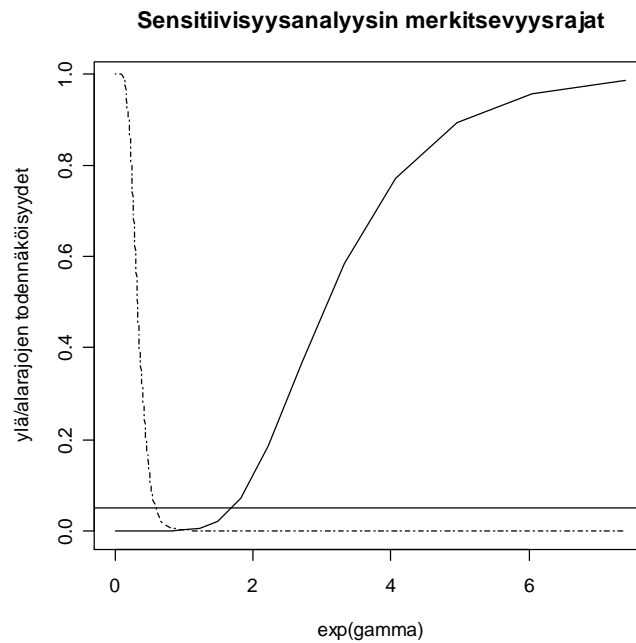
$$P(T^+ \geq a) \geq P(T \geq a | m) \geq P(T^- \geq a).$$

Sensitiivisyysanalyysissä merkitsevyydelle lasketaan useita rajoja eri g :n arvoilla. Yläraja $P(T^+ \geq a)$ jakautuu kuten $t(Z, y)$ kun $u_{si} = c_{si}$, ja alarajan $P(T^- \geq a)$ jakauma vastaa $t(Z, y)$:n jakaumaa kun $u_{si} = 1 - c_{si}$. Saatujen jakaumien perusteella merkitsevyydelle lasketaan todennäköinen väli. Jos analyysissä on saatu merkitsevä tulos, voidaan näin arvioida milloin tulokset saattaisivat olla ei-merkitseviä ja siten saada arvio tulosten uskottavuudelle.

Jos $P(T \geq a | m)$ noudattaa binomijakaumaa (kuten McNemarin testin tapauksessa), voidaan se rajoittaa välille

$$\sum_{a=t}^b \binom{b}{a} (p^+)^a (1-p^+)^{b-a} \geq P(T \geq t | m) \geq \sum_{a=t}^b \binom{b}{a} (p^-)^a (1-p^-)^{b-a},$$

jossa t on niiden parien määrä, joille $y_{Ti} = 1$ ja $y_{Ci} = 0$, ja b on epäyhtenevien parien kokonaismäärä. Jos testin tulos on ollut merkitsevä, saadaan eri g :n arvoilla ylä- ja alarajojen funktioiden kautta laskettua testin merkitsevyydelle väli. Tämä on esitetty kuviossa 1, kun oletetaan että $t = 30$ ja $b = 39$.



Kuvio 1. Sensitiivisyysanalyysin ylä- ja alarajojen funktiot.

Tällöin merkitsevyyden yläraja leikkaa suoran $P(T \geq a | m) = 0,050$ (eli merkitsevän rajan) välillä $1 < e^g = \Gamma < 2$. Esimerkin testi olisi sensitiivinen: piilevän harhan olisi lisättävä käsittelyyn joutumisen vedonlyöntisuhdetta melko vähän ($e^g < 2$), jotta testin tulokset muuttuisivat. Tulkinnat saattaisivat näin ollen muuttua olennaisesti jos piilevät kovariaatit huomioitaisiin.

5. Sovellus: Aktivoiko aktivointi?

Sovelluksessa pyritään vastaamaan kysymykseen näkykö työttömien aktivointisuunnitelmien laadinta heidän työllistymisessään. Vertailuryhmät kaltaistetaan rekisteristä saatavilla olevien muuttujien suhteen. Piilevän harhan vaikutusta tutkimustuloksiin pyritään havainnollistamaan sensitiivisyysanalyysillä.

5.1 *Aineisto ja asetelma*

5.1.1 URA -rekisteri

Rekisteritiedon käyttö liittyy hallinnollisiin päätöksiin, ja tieto on kerätty tätä silmällä pitäen. Koska tämän työn tulokset pohjautuvat rekisteriaineistoon, on tälläkin selvityksellä rekisteriaineistosta johtuen rajoitteensa. Aineisto on muodostettu työministeriön hallinnollisesta URA- rekisteristä. Se on luotu työvoimaviranomaisten tarpeiden pohjalta, ja siihen päivitetään asiakkaiden muutostiedot. URA -tietokanta on kattava, sillä siihen kirjautuvat kaikki Suomessa työvoimatoimistoihin työttömiksi ilmoittautuneet. Näin ollen siellä on myös kaikkien kirjattujen aktivointisuunnitelmien tiedot. Ainoastaan kuukauden 8., 18. ja 28. päivinä syntyneiden henkilöiden tiedot arkistoidaan, joten heiltä on tallessa myös kaikki aikaisemmat tapahtumat. Muiden osalta vanha tieto korvataan uudella tiedolla. Näin on esimerkiksi asuinpaikan ja aktivointisuunnitelman kohdalla. Joitakin selvästi virheellisiä tietoja rekisteriaineistosta myös löytyi. Nämä johtuivat ilmeisesti virhelyönneistä tai huolimattomuudesta. Muutaman havainnon jouduimme täysin poistamaan virheiden johdosta, mutta nämä tiedot eivät olleet onneksi kriittisiä tutkimuskysymyksen kannalta.

5.1.2 Kohde, rajaus ja työllistymisen määrittely

Vaikuttavuustutkimuksessa työllistyminen nähdään prosessina, ja sen tutkiminen sisältääkin aikaperspektiivin. Sen tutkiminen voi olla haasteellista. Tässä tutkielmassa mitataan ainoastaan aktivointisuunnitelmien laadinnan vaikutuksia työllistymiseen. Vasteessa ei näin ollen huomioitu kuntouttavan työtoiminnan lain sosiaalipoliittisia pyrkimyksiä, vaan keskityttiin laissa kirjattuun työllistämistavoitteeseen. Aktivointisuunnitelmat luovat olettavasti paremmat edellytykset työllistymiselle, vaikka työllistymisen syihin lukeutuvat myös muun muassa sattuma ja sosiaaliset suhteet, tai yleisemmin epäviralliset instituutiot. Rajauksen voi perustella määrällisillä lähtökohdilla, pyrkimyksellä löytää jotain uutta aikaisempiin tutkimuksiin verrattuna ja rekisteristä saatavan tiedon rajoituksilla. Tutkielman laajuus asetti osaltaan omat rajansa aineiston käsittelyn mahdollisuuksille.

Tutkimuspohjan muodostavat aktivointiehdon täyttävät, eli aktivointisuunnitelman laatimiseen oikeutetut työttömät. Koeryhmä aktivoidut sekä kontrolliryhmä ei-aktivoidut ovat otos tutkimuspohjan muodostavista yksilöistä. Aktivoidujen ryhmälle laadittiin aktivointisuunnitelma ja he osallistuivat suunnitelman mukaisiin toimiin, vertailuryhmä taas ei. Osa aktivoiduista ei välttämättä käytännössä osallistu toimiin jotka suunnitelmaan on listattu, mutta voidaan ajatella että pääosin suunnitelmat toteutuvat ja että jo suunnitelman laadinnalla on aktivoiva vaikutus. Esimerkiksi AKKU -tutkimuksessa jopa 83 % vastaajista koki jo aktivointisuunnitelman laadinnan olleen tarpeellinen tai erittäin tarpeellinen (Ala-Kauhaluoma ym. 2004: 125).

Suomessa aktivointitoimiin osallistumisen jaksot ovat verrattain pitkäkestoisia: vain kolmannes kestää 2 kuukautta tai alle, ja erilaisia jaksoja saattaa olla useita samalla henkilöllä. Ei olekaan mielekäästä puhua vertailuryhmästä joka ei olisi milloinkaan osallistunut mihinkään toimiin, sillä enemmistö suomalaisista pitkä-aikaistyöttömistä on osallistunut niihin jossain vaiheessa. (Aho ja Kunttu 2001: 7.) Vertailuryhmän määrittely helpottuu, kun vertaillaan ainoastaan aktivointisuunnitelman vaikutusta, ja oletetaan että muiden toimien vaikutus jakautuu tasaisesti eri tarkasteluryhmien välille. Tavoitteena onkin mitata aktivointisuunnitelmien työllistämisaikutusta muiden interventioiden lisänä.

Laki kuntouttavasta työtoiminnasta tuli voimaan 1.9.2001, mutta aluksi lain toteutus oli haparoivaa ja Suomen eri maantieteelliset alueet erosivat suuresti sen mukaan kenelle ja kuinka paljon

aktivointisuunnitelmia tehtiin. Tarkasteluikkunaksi rajattiin ajanjakso 1.1.2003–31.12.2005, sillä käytännöt vakiintuivat kunnissa riittävästi vuoteen 2003 tultaessa ja sittemmin vuoden 2006 alusta tuli voimaan niin kutsuttu työmarkkinauudistus, joka vaikutti siihen keneen aikaisempaa lakia sovelletaan, muuttaen sen velvoittavammaksi. Aineistossamme on kuitenkin tiedot jo lain voimaantulosta lähtien, joten pystyimme huomioimaan myös henkilöiden aikaisemman työllisyshistorian.

Työllistyminen määriteltiin tämän tutkielman puitteissa vuosittaiseksi tapahtumaksi. Henkilön katsottiin työllistyneen, jos hän oli työssä avoimilla työmarkkinoilla vähintään 90 päivää vuodessa. Valittua 90 päivän kriteeriä kutsutaan työllistymisehdoksi. Valittu työllistymisehto on verrattavissa aikaisemmissa tutkimuksissa käytettyihin kriteereihin. Esimerkiksi Aho ja Kunttu (2001) käyttivät kolmen kuukauden rajaa selvittäessään toimien vaikuttavuutta työllistymiseen. Myöhemmin Aho (2004) määritteli tutkimuksessaan kroonisen työttömyyden jatkuvan, vaikka henkilö olisi ollut työssä 10 prosenttia työvoimaankuulumisajasta, jolloin työ on vain satunnaista, eikä voida puhua varsinaisesta työllistymisestä. Vuodesta 10 prosenttia on noin 37 päivää, joka on yli puolet (59 %) lyhyempi ajanjakso kuin valitsemamme 90 päivää. Lyhyet työjaksot avoimilla työmarkkinoilla olivat aineistossamme kuitenkin suhteellisen yleisiä, ja jopa neljäsosa kaikista työpätkistä kesti 3 päivää tai alle. Tämä voi johtua siitä, että monet työttömät hakivat työtä mm. palvelu-, toimisto- tai teollisuusosalta (yht. 34,4 %), joissa työnvälityksen kautta esimerkiksi kiireapulaisen tehtävät ovat kestoaltaan lyhyitä, ja työt pätkittäisiä. Kun kriteeriksi valitaan 90 päivää, eivät useat hyvin lyhyet työssäolajakso vielä täytä työllistymisehtoa, vaan henkilön on oltava erityisen aktiivinen työnhaussa, tai oltava pidempiä jaksoja työssä jotta työllistymisehto täytyisi.

Aineistossa oli tiedot kaikista työllistymiseen liittyvistä tapahtumista, joista rajasimme osan koskemaan tässä käytettävää työllistymisen määritelmää. Jotta päivä laskettiin työllistymisehdon mukaiseksi työssäolopäiväksi, tutkittavan tuli olla työllistynyt avoimille työmarkkinoille, ja yksikin päivä riitti kartuttamaan työssäolon kokonaisaika. Määritelmässä ei huomioitu oliko työsaanti tapahtunut itsenäisesti vai työvoimatoimiston avustuksella.

5.1.3 Otanta

Otanta oli kaksiasteinen. Ensin valittiin Suomen viidestätoista työvoima- ja elinkeinokeskuksesta (TE -keskus) mukaan neljä aktivoinnin ja työllistymisen suhteen erilaista aluetta. Näin ollen otos on ryvästynyt TE -keskuksittain. Toiseksi alueilta valittiin mukaan vain ne aktivointiin oikeutetut, joiden tiedot oli tallennettu (suurelta osin) pysyvästi tietokantaan. Tämä voidaan tulkita satunnaisotannaksi. Kaksiasteisen otannan jälkeen oli aineistossa tiedot 44622 henkilöltä.

Alueiksi valittiin TE -keskukset, sillä ne ovat riittävän suuria, jotta yksittäisten työvoimatoimistojen linjausten vaikutukset esimerkiksi aktivoitavien valikoinnissa saatiin minimoitua. TE -keskus alueilta oli myös saatavilla luotettavaa tietoa vuosittaisista työllisyyden muutoksista mm. työnhakualoittain. Valitut neljä aluetta ovat Etelä-Savo, Lappi, Uusimaa ja Pohjanmaa.

Aluevalinnan maantieteellinen perustelu juontaa juurensa siitä, että näillä alueilla on erilaiset aktivointiasteet, työnsaantimahdollisuudet ja työttömyyskehitys (Aho ja Kunttu 2001). Aktivointiaste on niiden henkilöiden lukumäärä joille on tehty aktivointisuunnitelma suhteutettuna aktivointiehtdon täyttävien henkilöiden määrään. Tämä tieto oli saatavilla ajalta 1.9.2001–23.2.2004. Taulukossa 1 näkyvät valittujen TE -keskusalueiden työllisyys- ja aktivointiasteet keskimäärin ajalta 2001 - 2004. (Ala-Kauhaluoma ym. 2004; TE -keskukset)

Taulukko 1. Tutkittavien TE -keskusten keskimääräiset aktivointi- ja työllisyysasteet.

TE-keskus	aktivointiaste työllisyysaste	
Etelä-Savo	17 %	62 %
Lappi	81 %	59 %
Pohjanmaa	75 %	69 %
Uusimaa	55 %	74 %

Taulukosta nähdään, että Etelä-Savossa aktivointiaste on ollut erityisen matala. Työllisyysaste on siellä keskimääräinen (62 %). Lappi ja Pohjanmaa ovat korkean aktivointiasteen alueita, ja niissä on tehty aktivointisuunnitelmia yli ¾ :lle kaikista oikeutetuista. Lapissa työllisyysaste on kuitenkin matala (alle 60 %), ja työttömyys korkea. Pohjanmaalla taas työllisyysaste on keskimääräinen, ja työttömiä on vähemmän kuin muualla Suomessa keskimäärin. Uusimaa eroaa muista korkealla

työllisyysasteellaan. Pääkaupunkiseudulla onkin hyvät työnsaantimahdollisuudet. Aktivointiaste on siellä jäänyt keskimääräisen 55 prosentin tuntumaan.

5.1.4 Asetelma

Mahdollisia asetelmavaihtoehtoja on karkeasti esiteltynä taulukossa 2. Ajatuksena on erotella tietyllä aikavälillä, seurantavuoden aikana ja jatkuvassa seurannassa tapahtuva työllistymisen tarkastelu.

Taulukko 2. Erilaisia asetelmavaihtoehtoja.

AIKAVÄLI	tietyin ajan kuluttua tietyin ajan kuluessa	työllistyi / ei työllistynyt työllistyi / ei työllistynyt
SEURANTAVUOSI	seurantavuoden lopussa seurantavuoden aikana kuinka kauan ollut työssä seurantavuonna	työllistyi / ei työllistynyt työllistyi / ei työllistynyt 0-12 (kk)
JATKUVA SEURANTA	kuinka nopeasti sai töitä kuinka kauan ollut työssä tiettyyn hetkeen mennessä	0-> (kk) 0-> (kk)

Työllistyminen määriteltiin vuosittaisena tapahtumana, joten jatkuva seuranta ei ollut mahdollinen asetelma. Tietylle aikavälille rajoittuvassa asetelmassa seurattaisiin työllistymistä aktivointisuunnitelman laadinnan päivämäärästä eteenpäin esimerkiksi kahden vuoden ajan. Tämä ei ollut mahdollista aineiston rajoitusten vuoksi, vaikka se kuvaisi työllistymiskehitystä parhaiten näistä asetelmavaihtoehdoista. Jos työllistymistä vertailtaisiin ryhmissä aikavälin tai seurantavuoden lopussa, saisimme vain yhden hetken (tässä päivän) tiedon, joka ei taas vastaa työllistymiskehitystä ja sen määritelmää.

Päädyimme asetelmaan, jossa seurataan yksilöiden työllistymisehdon täyttymistä seurantavuosina. Perusvuosi on asetelmassa 2003, jonka aikana koeryhmään valikoituville laadittiin aktivointisuunnitelma. Koeryhmälle haimme kaltaistetun vertailuryhmän aktivointiin oikeutetuista henkilöistä, joille ei ollut koskaan laadittu aktivointisuunnitelmaa. Käsittelyryhmät valikoituivat

näin ollen erilaisista aktivointiin oikeutettujen joukoista. Seurantavuosina 2004 ja 2005 tutkittiin koe- ja kontrolliryhmän työllistymistä.

Asetelmavalinta vastaa tapaus-kohortti -lähestymistapaa. Se on kohorttitutkimus, jossa valittua tutkimusryhmää seurataan tietty aika, ja johon on yhdistetty tapaus-verrokki-asetelma koe- ja verrokkiryhmineen (Clayton ja Hills 1993: 329-330). Kuviossa 2 on esitetty aktivoitujen ja ei-aktivoitujen seurantaa. Kuviossa 3 näkyy asetelma, jossa jokaiselle aktivoidulle on etsitty sekoittavien tekijöiden suhteen kaltaistettu aktiivimaton verrokki. Kuvioissa A tarkoittaa aktivointiajanhetkeä, ja T työllistymistä.

aktivointi	2003	2004	2005
aktivoidut	A-----T		
	A-----		
	A----T		
	A-----		
ei -aktivoidut	-----T		
	-----T		

	-----T		
	-----T		

Kuvio 2. Aktivoitujen ja ei-aktivoitujen seurantaa.

vastinparit	2003	2004	2005
aktivoitu	A-----T		
ei -aktivoitu	-----T		
aktivoitu	A-----		
ei -aktivoitu	-----T		
aktivoitu	A----T		
ei -aktivoitu	-----T		
aktivoitu	A-----		
ei -aktivoitu	-----		

Kuvio 3. Vastinparien seurantaa kaltaistetussa asetelmassa.

Aktivointisuunnitelmiin kirjatut toimet ajoittuivat 1.1.2003 jälkeiselle ajalle, ja ne saattoivat jatkaa seurantavuosien aikana. Tässä tutkielmassa oletettiin, että aktivointisuunnitelman laadinnan vaikutus kesti seurantavuosien ajan, ja että tämä on riittävä aika saada näkymään sen mahdollinen vaikutus työllistymiseen. Aikaisemmissa tutkimuksissa on jo selvitetty, että jo yhden vuoden seurannalla saadaan hyvin samantapaiset tulokset kuin pidemmällä aikavälillä (esim. Aho ja Kunttu 2001).

Aineistossa 295:lle aktivointisuunnitelma oli tehty vuoden 2003 aikana. Koeryhmään valikoitui näin ollen 17 prosenttia kaikista joille oli tutkimusaikana tehty aktivointisuunnitelma. Koeryhmä vastasi taustamuuttujiltaan muina aikoina aktivoituja. Muulloin aktivoiduissa Lapin ja Etelä-Savon TE -keskusten osuudet korostuivat, kun taas Uudeltamaalta oli suhteessa vähemmän havaintoja vuonna 2003 aktivoituihin verrattuna. Erot olivat kuitenkin pieniä, joten koeryhmän voidaan olettaa kuvaavan hyvin aktivoituja yleensä. Asetelmavalinta on esitetty tiivistetysti alla.

Perusvuosi 2003, jonka aikana aktivoidut muodostavat koeryhmän.
Kontrolliryhmälle ei ole laadittu aktivointisuunnitelmaa koko tutkimusaikana.
Työllistymistä vertaillaan vuosina 2004 ja 2005 työllistymisehdon perusteella.

5.2 Aineiston kuvailua

5.2.1 Muuttujat

Vaikutussuhde on tässä tutkielmassa yksinkertaistettu operationalisoinnin avulla. Rajausten kautta tutkittiin vain tiettyjen rekisteristä löytyvien, sisällöllisesti olennaisten ja vaihteluun vaikuttavien muuttujien vaikutusta työllistymiseen. Muuttujat on luetteloitu taulukossa 3.

Taulukko 3. Aineiston muuttujat.

Henkilötiedot	Ikä	vuosissa
	Luokiteltu ikä	18-24, 25-34, 35-44, 45-54, 55-64, 65-
	Sukupuoli	0 Nainen
		1 Mies
	Koulutustaso	1 Perusaste
		2 Keskiaste
		3 Korkea-aste
	TE -keskusalue	1 Etelä-Savo
2 Lappi		
3 Pohjanmaa		
4 Uusimaa		
Vajaakuntoisuus	1 Vajaakuntoinen	
	0 Työkykyinen	
Työnhakuala	1 Valkokaulus (terveys, hallinto jne)	
	2 Perus (kuljetus, rakennus, teollisuus jne)	
	3 Palvelu tai muu	
Työllisyystiedot	Aktivointi	1 aktivoitu
		0 ei aktivoitu (koko tutkimusaikana)
	Aktivointipäivämäärä	pp.kk.vvvv
	Kurssi	0 Ei kurssia
		1 Kurssi, aloitettu ennen 1.2003
	Sijoitus	0 Ei ollut sijoitettuna
		1 Sijoitettuna, ennen 1.2003
	Työssäolopäivät (vuosittain 01-05)	Työssäolopäivien lkm
Työllistyminen (vuosittain 01-05)	0 Ei työllistynyt	
	1 Työllistyi	
Työllistymisten lkm	Ajalla 9.2001 - 1.2006	

Käsittelytekijänä on aktivointisuunnitelman laatiminen. Suunnitelman allekirjoituspäivämäärä oli välillä 1.9.2001-31.12.2005 1710 henkilöllä (3,8 %) kaikista aineiston työttömistä (44622 henkilöä). Allekirjoituspäivämäärä vastaa tässä tutkielmassa aktivointipäivämäärää. Käsitelimme ainoastaan suunnitelman allekirjoittaneita, sillä muilta ei ole tietoa aktivointisuunnitelman laadinnan ajanhetkestä. Loput, eli henkilöt joille suunnitelma on tehty mutta jotka eivät allekirjoittaneet sitä, 0,05 prosenttia (22 henkilöä) kaikista työttömistä, olivat samankaltaisia muiden mitattujen kovariaattien suhteen verrattuna allekirjoittaneiden joukkoon.

Aktivointisuunnitelmätieto oli kuitenkin harhaista, sillä uuden suunnitelman laadinta korvasi entisen, jolloin aineistossa näkyivät vain myöhemmät suunnitelmat. Näin ollen ehkä kaikkein vaikeimmin työllistyvät, joille suunnitelmaa on päivitetty myöhemmin, eivät valikoituneet koeryhmään. Nuorten, alle 25-vuotiaiden, osalta aktivointitieto on luotettavinta, sillä useille heistä suunnitelmia on tehty vain yksi. Mitä kauemmin ollut työttömänä, ja näin ollen myös mitä iäkkäämpi työtön henkilö on, sitä herkemmin monissa työvoimatoimistoissa uusittiin suunnitelmia. Toisaalta yksi priorisointi aktivointisuunnitelmien laadinnalle oli mahdollisimman varhainen puuttuminen, jolloin nuorille tehtiin ehkä herkästi ensimmäinen aktivointisuunnitelma. Käytännöt vaihtelevat tosin suuresti työvoimatoimistoittain ja työntekijöittäin.

Ikä laskettiin syntymäajasta tutkimushetken 1.1.2003. Tämän perusteella se luokiteltiin 5 luokkaan. Alin tutkittu ikäluokka kattaa 18–24 -vuotiaat. Seuraavat luokat on jaettu 10 vuoden välein.

Dikotomisia muuttujia olivat sukupuoli, vajaakuntoisuus, kurssi ja sijoitus. Vajaakuntoisuus kuvaa henkilön saaneen diagnoosin, jonka perusteella hänet voidaan luokitella joko psykologisin tai fysiologisin perustein vaikeasti työllistyväksi. Näitä syitä ei lähdetty erottelemaan, vaan käsiteltiin kaikkia luokituksen saaneita yhtäläillä vajakuntoisina työllistymismahdollisuuksiin nähden. Vajaakuntoisuudelle oli tiedossa myös toteamispäivämäärä. Luokittelimme vajaakuntoisiksi kuitenkin kaikki, jotka saivat jollain perusteella vajaakuntoisuusluokituksen tarkasteluaikeavälillä. Tämä johtui siitä, että esimerkiksi aktivoinnin myötä "löydettiin" työnhakijoiden joukosta todellisuudessa vajaakuntoisia työnhakijoita, jotka siis olivat vajaakuntoisia jo ennen toteamispäivämäärää. On mahdotonta erotella vajaakuntoisista niitä, joilla päivämäärätieto on perusteltu, joten tutkielmassa käytettiin vain karkeaa tietoa siitä, että henkilö todettiin vajaakuntoiseksi jossain vaiheessa tutkimusaikana.

Kurssi -muuttuja sai arvon 1, jos henkilö oli osallistunut kurssille (työvoimapolitiisena toimenpiteenä), ja jos kurssi oli alkanut ennen vuotta 2003. Sijoitus -muuttujasta muodostettiin myös dikotominen, jolloin se sai positiivisen arvon jos henkilö oli ollut sijoitettuna ennen vuotta 2003 mille tahansa sektorille. Koulutustieto taas luokiteltiin karkeasti perusasteen, keskiasteen ja korkea-asteen koulutukseen. Tutkittavilta henkilöiltä oli tiedossa näiden lisäksi myös työllistymisten, eli työn alkamispäivien, lukumäärät.

Työnhakuala on rekisteriin merkitty ala jolta henkilö hakee työtä ensisijaisesti. Aloja oli useita, ja ne luokiteltiin kolmeen. Luokittelu on esitetty taulukossa 4.

Taulukko 4. Työnhakualat.

Toimisto	Tieteellinen, tekninen, taiteellinen, humanistinen, terveys, sosiaali, hallinto, toimisto, it, kauppa
Teollisuus	Maa, metsä, kala, kuljetus, liikenne, rakennus, kaivos, teollisuus, varasto
Palvelu tai muu	Palvelu, tuntematon

Aineistosta löytyi selkeitä virheitä useista muuttujista. Näiden henkilöiden tiedot jouduttiin jättämään analyysin ulkopuolelle. Muiden kovariaattien osalta poistetut vastasivat keskimääräisiä tietoja aineistossa.

5.2.2 Alustavat tarkastelut

Aktivointiasteet erosivat aineistossa oletetusti TE -keskuksittain. Etelä-Savossa aktivoituja oli suhteessa huomattavasti vähemmän kuin muilla alueilla. Taustamuuttujiltaan aktivoituneet ja aktivoimattomat erosivat erityisesti sukupuolen ja vajaakuntoisuuden osalta siten, että aktivoitujen joukossa oli enemmän miehiä ja vajaakuntoisia. Vertailujoukot erosivat myös koulutuksen suhteen: korkeasti koulutettuja ei aktivoitu niin ahkerasti. Työnhakualoittain aktivoituja oli eniten luokittelemattomilla, sekä teollisuus- ja varastoaloilla. Vertailtaessa ryhmien työllistymistä kontrolliryhmä työllistyi paremmin kuin koeryhmä, vaikka koeryhmästä puuttuivatkin mahdollisesti kaikkein vaikeimmin työllistyvät, joille aktivointisuunnitelma oli uusittu myöhemmin perusvuoden 2003 jälkeen.

Aineisto koostui aktivointiin oikeutetuista neljältä eri työvoima- ja elinkeinokeskus – alueelta. Uudenmaan alue painottuu selvästi (63,6 % havainnoista), kun havaintoja on muilta alueilta jokaisesta vain hieman yli 10 prosenttia. Aktivointisuunnitelmia laadittiin alueittain epätasapainoisesti, ja aktivoitujen joukossa Uudenmaan osuus oli hieman tasoittunut, vaikka sieltä oli edelleen yli puolet (55,4 %) havainnoista. Lapista oli aktivoituista 25,6 prosenttia.

Taulukossa 5 näkyy aineistosta lasketut aktivoitujen osuudet (sulkeissa vuoden 2003 osuudet) alueilla kaikista aktivointiin oikeutetuista. Etelä-Savo erottuu muista alueista vähäisellä aktivoitujen määrällä (vain 1,7 %, vuonna 2003 0,1 % aktivointiin oikeutetuista). Kokonaisuudessaan aktivointisuunnitelman allekirjoittaneita oli tutkimusaikavälillä aineistossa 3,8 % kaikista aktivointiin oikeutetuista.

Taulukko 5. TE- keskus ja aktivoitujen osuudet. (Suluissa vuoden 2003 osuudet)

TE -keskus	aktivoituja
Etelä-Savo	1,7 % (0,1 %)
Lappi	7,1 % (1,0 %)
Pohjanmaa	4,5 % (0,7 %)
Uusimaa	3,4 % (0,7 %)

Aineistossa (44622 henkilöä) oli vajaakuntoisia 9,3 %, miehiä 49,0 % ja keskiasteen koulutuksen käyneitä 43,8 %. Aktivoiduissa (1710 henkilöä) oli suhteessa runsaasti vajaakuntoisia 27 % ja miehiä 62,5 %. Aktivoidut ja aktivoimattomat erosivatkin erityisesti sukupuolen ja vajaakuntoisuuden osalta. Vajaakuntoisista jopa 11 %:a oli aktivoitu, kun ei-vajaakuntoisista sama osuus oli vain 3 %. Tämä saattaa selittyä myös sillä, että aktivointiprosessissa aktivointiin oikeutettujen pitkäaikaistyöttömien joukosta osa sai aktivoinnin myötä myös ko. luokituksen.

Keskiasteen koulutuksen käyneitä oli aktivoiduissa 39,4 %. Aktivoiduissa oli myös runsaasti perusasteen koulutuksen käyneitä. Mitä korkeampi koulutus henkilöillä oli, sitä suurempaa osaa ei koskaan aktivoitu. Työnhakualoittain taas aktivoiduilla painottuivat aktivoimattomia enemmän luokittelemattomat alat (25,7 %) sekä teollisuus ja varastoalat (yhteensä 23,7 %). Kokonaisuudessaan toimihenkilöaloille heistä pyrki yhteensä vähemmän (24,9 %) kuin aktivoimattomista (43,7 %).

Vuonna 2003 aktivoiduista (295 henkilöä) 18 % oli ollut sijoitettuna, kun ei –aktivoiduista sijoitettuna oli ollut vain 7 %. Kurssien osalta prosentuaaliset osuudet olivat aktivoiduille 11 % ja aktivoimattomille 5 %. Iältään aktivoitujen lukumäärät painottuvat luokkiin 35–44 ja 45–54 -

vuotiaat (molemmissa yli 24 % aktivoituista), kun kontrolliryhmän ikä on jakautunut tasaisesti viiteen luokkaan. Voimakkaasti yleistäen aktivoitunut ovat vähemmän koulutettuja erityisen vaikeasti työllistyviä miehiä, jotka hakevat työtä tuotantoaloilta.

Suurin osa työllistymisjaksoista oli lyhytkestoisia, ja aineistossa oli henkilöitä jotka olivat käytännössä olleen työssä ja työttömänä lähes vuoropäivin. Aineistosta voitiin selvittää myös jaksojen pituudet. Työssäoloaikojen mediaani oli 19 päivää, ja työttömyysjaksojen taas 44 päivää. Työ- ja työttömyysjaksot olivat jakautuneet oikealle vinosti, joten lyhyitä jaksoja oli runsaasti, ja pidemmät (vuosien mittaiset) jaksot olivat harvinaisia. Erityisesti työttömyysjaksoissa löytyi (5 %) pitkäkestoisia, yli 400 päivää kestäviä, yhtenäisiä jaksoja.

Aktivoitujen joukossa työllistyminen oli selvästi heikompaa kuin aktivoimattomien joukossa (42890 henkilöä). Taulukossa 6 on esitetty eri vuosina aktivoitujen työllistyminen (työllistymisehdon mukaan) aktivointia seuraavina vuosina. Tästä näkyy jo alustava tulos aktivoinnin työllistämisaikutuksesta.

Taulukko 6. Aktivoitunut ja työllistyminen.

Vertailuryhmät	työssä v.2003	työssä v.2004	työssä v.2005
ei-aktivoitunut	16,2 %	16,3 %	14,2 %
aktivoitu 2002	9,5 %	9,1 %	6,1 %
aktivoitu 2003	-	6,8 %	6,8 %
aktivoitu 2004	-	-	3,5 %

Eri vuosina aktivoituista oli seuraavana vuonna työllistynyt 3,5 - 9,5 prosenttia, kun ei-aktivoituista työllistyi koko tarkasteluajanjaksolla vuosittain keskimäärin 15,6 prosenttia. Kokonaisuudessaan työllistyneitä oli tutkimusjoukossa vähän. Aktivoitujen seuraavan vuoden työllistymisluvut vähenevät vuosittain niin, että vuonna 2004 aktivoituista enää vain 3,5 prosenttia työllistyi vuonna 2005. Tämä saattaa selittyä sillä, jos aktivoituiksi valikoitui ajan myötä yhä vaikeammin työllistyviä. Kontrolliryhmä työllistyi huomattavasti paremmin, mutta ryhmät eroavat vielä tässä vaiheessa useiden muuttujien suhteen, jotka voivat sekoittaa tulosta.

Kun tarkastellaan vuoden 2004 työllistymistä, koeryhmässä perusasteen koulutuksen saaneet olivat erityisen huonosti työllistyneitä, kuten myös jos henkilö oli vajaakuntoinen, tai nainen. Havaintoja

oli tosin vähän, joten arviot ovat vain suuntaa antavia. Iältään 25-34 –vuotiaat henkilöt olivat työllistyneet suhteellisesti hyvin (21,3 %). Kontrolliryhmästä taas kurssin käyneet työllistyivät hieman paremmin, ja palvelualoille hakeutuvat huonommin kuin muut. Mielenkiintoista oli, että sekä koe- että kontrolliryhmässä suurempi osa niistä henkilöistä jotka eivät olleet työllistyneet vuonna 2002, työllistyivät vuonna 2004. Aho ja Koponen (2007) havaitsivat tutkimuksessaan, että työllistymisen suhteen eroja löytyi iän, koulutustason, vajaakuntoisuuden, työmarkkinahistorian sekä TE -keskusten suhteen. Sukupuoli ei erotellut työnhakijoita. Samansuuntaisia tuloksia saatiin tässäkin tutkielmassa, erot tosin vaihtelivat käsittelyryhmien välillä. Taustamuuttujien frekvenssit ja suhteelliset osuudet aktivoinnin ja vuoden 2004 työllistymisen suhteen löytyvät liitteestä 2.

5.3 Harhan kontrollointi ja käsittelyvaikutus

5.3.1 Harha

Vaikutussuhde ei ole yhteiskunnallisissa ongelmanasetteluissa koskaan puhdas, vaan asetelmaan vaikuttaa useita sekoittavia tekijöitä. Jos kontrolloidaan vaikuttavan tekijän seurauksia, tuo tämä harhaa päättelyyn (Ho ym. 2006: 3). Ongelma onkin miten huomioida kokonaisvaltaisesti asiat, jotka vaikuttavat tutkittavaan syy-seuraussuhteeseen. Sekoittavien tekijöiden vaikutus on tässä työssä pyritty kontrolloimaan kaltaistuksella. Otantaa voitaneen pitää satunnaisotantaan verrattavana, joten oli mahdollista käyttää perinteisiä menetelmiä käsittelyvaikutuksen estimoinnissa.

Puuttuvat havainnot eivät yleisesti ottaen ole suuri ongelma rekisteritutkimuksissa. Rekisteritiedossakin saattaa olla kuitenkin puutteita, ja toisinaan imputoidaan puuttuvien havaintojen kohdalle tietoa esim. muiden yksilöiden tietojen avulla. Sensurointi tarkoittaa kesken tutkimusta kadonneita yksilöitä, eli henkilön kesken jäänyttä seurantaa, jolloin ei voida tietää vasteen lopullista arvoa (Clayton ja Hills, 1993: 24–25). Voidaan ajatella, että tutkielmassa osa henkilöistä sensuroitui puutteellisen tiedonsaannin johdosta. Emme esimerkiksi pystyneet varmasti identifioimaan kaikkien henkilöiden työllistymistä, sillä aineiston yksi puute oli, että joissakin tapauksissa henkilöiden tiedot päättyivät "ei ole uusiutunut työnhakuaan" -merkintään. Tällöin he

sensuroituivat, vaikka todellisuudessa osa heistä saattoi olla myös työllistyneitä, mutta he eivät ilmoittaneet sitä työvoimatoimistoon. Tiedämme vain milloin henkilö on poistunut työttömyyskortistosta, mutta emme aina syytä. Tätä varten olisi pitänyt yhdistellä tietoja eri rekistereistä. Tätä ongelmaa ovat selvittäneet Aho ja Koponen (2007) seurannassa, jossa todetaan työttömien osuuksien alenevan, kun seuranta-aikaa pidennetään. Tässä tutkielmassa vasteen arvoihin liittyy harhaa, jonka suuruutta ei pystytä arvioimaan tällä aineistolla.

Havaittavissa olevaa harhaa voitiin kontrolloida mitatuilla sekoittavilla tekijöillä, joiden vaikutus minimoitiin käyttämällä niitä kaltaistavina muuttujina. Piilevää harhaa saattoi tulla asetelmaan esimerkiksi aktivointisuunnitelman laadinta-ajankohdasta, josta ei ollut mahdollista saada täsmällistä tietoa. Tutkimuksen ulkopuolelle jäi myös havaitsemattomia tekijöitä, kuten henkilön työnhakumotivaatio ja sosiaaliset taidot (Aho ja Kunttu 2001: 5) sekä työvoimatoimistojen linjausten vaikutus. Linjausvaikutus pyrittiin kuitenkin huomioimaan valitsemalla useita työvoimatoimistoja kattavia TE -keskus -alueita mukaan tutkimukseen. Ei-mitatut tekijät saattavat vaikuttaa työllistymistodennäköisyyteen. Sensitiivisyysanalyysillä voidaan selvittää kuinka herkkiä tulokset ovat havaitsemattomien tekijöiden vaikutukselle.

5.3.2 Kaltaistus

Yksinkertainen lähestymistapa olisi ollut kaltaistaa yksilöt suoraan havaittujen kovariaattien suhteen siten, että kaltaisilla pareilla olisi samat kovariaattien arvot, kuten Aho ja Kunttu (2001). He joustivat kaltaistuksen tarkkuudesta, jotta saivat enemmän (11 %) havaintopareja, jolloin suhteellisen pienillä joustoilla löydettiin lähes kaikille kaltaistettu pari. On kuitenkin joustavampaa käyttää kaltaistuksessa menetelmiä, jotka huomioivat aktivointitodennäköisyyden yksittäisten kovariaattien arvojen sijaan.

Aluksi kontrolliryhmää pienennettiin satunnaisotannalla siten, että ryhmään valittiin vain 8. päivä syntyneet. Alun perin ei-aktivoituja oli 42890 henkilöä, mutta otannan jälkeen joukko kutistui 14054:ään (32,8 %). Koeryhmä, eli vuonna 2003 aktivoitunut, tuli kokonaisuudessaan mukaan, sillä heitä oli aineistossa ainoastaan 295 henkilöä. Kaltaistuksessa koeryhmän yksilöille etsittiin verrokkit supistetusta kontrolliryhmästä.

Kaltaistavina muuttujina käytettiin sekoittavia tekijöitä, jotka siis vaikuttivat sekä aktivointiin, että binääriseen työllistymis-muuttujaan. Yhteys testattiin, ja tuloksena lähes kaikki muuttujat olivat sekoittavia tekijöitä. Loputkin muuttujat olivat sisällöllisesti olennaisia, joten niitä ei jätetty kaltaistuksesta pois. Taulukossa 7 on lueteltu kaltaistuksessa käytetyt muuttujat.

Taulukko 7. Kaltaistusmuuttujat.

Muuttujat
Ikä
Sukupuoli
Koulutus
TE -keskusalue
Vajaakuntoisuus
Työnhakuala
Kurssi -muuttuja
Sijoitus -muuttuja
Työssäolopäivien määrä vuonna 2001
Työssäolopäivien määrä vuonna 2002

Työttömien aikaisempien vuosien tapahtumat otettiin huomioon vain työssäolopäivinä. On mahdollista, että työssäolopäivien mukana saadaan mukaan myös taustalta tietoa esimerkiksi työnhakumotivaatiosta, sillä aiemmissa tutkimuksissa on todettu, että asiakkaan työnhaun aktiivisuus on yhteydessä työnsaantiin (mm. Ala-Kauhaluoma ym. 2004).

Aineisto olisi ollut mahdollista osittaa joidenkin näiden muuttujien suhteen, ja tutkia työllistymistä osajoukoittain. Esimerkiksi nuorten (alle 25-vuotiaiden) työllistymistä olisi ollut mielenkiintoista seurata erikseen, sillä heidän aktivointisuunnitelmiansa päivämäärätieto oli luotettavinta. Myös TE -keskusalueiden tutkiminen erikseen olisi perusteltua jo erilaisen työllisyys- ja aktivointikehityksen johdosta. Tämän tutkielman puitteissa keskityimme kuitenkin kaltaistuksen ja myöhemmin piilevän harhan tutkimiseen.

Tutkielmassa käytettiin kaltaistusperusteena propensiteettipistemäärää. Tällöin oletettiin, ettei piilevää harhaa ole, ja propensiteettipistemäärä voitiin estimoida aktivointitodennäköisyyksien $p(x) = P(Z = 1 | x)$ avulla. Aktivointitodennäköisyyksien estimoinnissa käytettiin logistista regressiota. TE – keskusalue ja työnhakuala käsiteltiin dummy –muuttujina, ja selitettävänä oli dikotominen käsittely (aktivoitu vuonna 2003 / ei-aktivoitu). Ennustetut todennäköisyydet kuulua käsittelyryhmään

$$\hat{x}(x) = \frac{e^{x'\hat{B}}}{1 + e^{x'\hat{B}}}$$

saatiin estimoitujen \hat{B} avulla, jotka löytyvät taulukosta 8 sarakkeesta B.

Taulukko 8. Logistinen regressio.

	B	S.E.	Wald	df	Sig.	Exp(B)
Muuttujat sukupuoli	-0,421	0,135	9,704	1	0,002	0,657
ikä	0,003	0,005	0,456	1	0,499	1,003
vajaakuntoisuus	1,336	0,139	93,086	1	0	3,806
sijoitus	0,839	0,164	26,217	1	0	2,315
kurssi	0,716	0,2	12,828	1	0	2,045
tyovuosi_01	-0,005	0,006	0,507	1	0,476	0,995
tyovuosi_02	-0,008	0,002	18,279	1	0	0,992
te1	-0,456	0,156	8,501	1	0,004	0,634
te2	-2,232	0,432	26,687	1	0	0,107
te3	-0,277	0,209	1,748	1	0,186	0,758
am1	-0,168	0,17	0,979	1	0,322	0,845
am2	0,05	0,16	0,1	1	0,752	1,052
koulutus	-0,536	0,099	29,109	1	0	0,585
Constant	-2,597	0,293	78,723	1	0	0,074

Taulukosta nähdään, että aikaisemmin suoritetulla sijoituksella tai kurssilla on positiivinen yhteys aktivointiin. Vajaakuntoisuuden ja aktivoinnin yhteys on erityisen voimakas. Tulokset myös muiden muuttujien osalta ovat samansuuntaisia kuten edellisissä alustavissa tarkasteluissa. Huomattavaa on, että aikaisemman työssäolon kerroin on lähes nolla, kuten iänkin.

Aktivointi oli harvinainen tapahtuma, joten malli ennusti kaikkien yksilöiden kuluvaan kontrolliryhmään. Ennustettujen todennäköisyyksien kuvaaja on oikealle vino (Liite 3), ja aktivointitodennäköisyydet olivat kaikille yksilöille lähellä nollaa: niiden keskiarvo oli 0,021. Koska todennäköisyydet olivat kaikki niin pieniä, ei aineistoa ositettu sen perusteella. Lähimmän naapurin menetelmässä kaltaisiksi valikoituvat ne yksilöt, jotka ovat lähimpänä toisiaan propensiteettipistemäärän perusteella. Kaltaistuksessa käytettiin R 2.4.0 -ohjelman ohjelmapakettia `matchit {MatchIt}`. Näin jokaiselle vuonna 2003 aktivoitulle löydettiin yksi kaltainen verrokki supistetusta kontrollijoukosta. Kaltaistuksen jälkeen kaikki sekoittavat muuttujat olivat tasapainossa käsittelyryhmissä, joten ryhmät olivat vertailukelpoisia. Tämän jälkeen oli mahdollista käyttää satunnaistamisen olettavia tilastollisia menetelmiä, sillä käsittelyryhmät eivät eronneet tilastollisesti havaittujen muuttujien suhteen. Havaitsemattomien muuttujien mahdollista vaikutusta arvioitiin tämän jälkeen.

5.3.3 Käsittelyvaikutuksen testaus

Käsittelyvaikutusta testattiin McNemarin testillä, joka soveltuu binääriselle pareittaiselle aineistolle. Aineisto muodostuu nyt $S = 295$ kaltaistetusta koe-verrokki-parista. Nollahypoteesina on, että vertailuryhmät eivät eroa, ja vastahypoteesi on kaksisuuntainen. Testisuure oli muotoa

$$T = t(Z, y) = \sum_{s=1}^S \sum_{i=1}^2 c_{si} Z_{si},$$

missä $Z_{si} = 1$ jos i :s henkilö parissa s on aktivoitu ja vastaavasti $Z_{si} = 0$ jos tämä henkilö kuuluu kontrolliryhmään, ja $c_{si} = 1$ jos $y_{si} = 1$ eli henkilö työllistyi, muutoin 0. Näin ollen testisuure ilmoittaa työllistyneiden aktivoitujen määrän. Seuraavan sivun taulukoissa näkyy aktivoitujen ja kaltaistettujen verrokkien työllistyminen vuosina 2004 ja 2005.

Taulukko 9. Työllistyminen vuonna 2004.

vuosi 2004		aktivoitunut		Yhteensä
		ei työllistynyt	työllistyi	
verrokit	ei työllistynyt	252	16	268
	työllistyi	23	4	27
Yhteensä		275	20	295

Taulukko 10. Työllistyminen vuonna 2005.

vuosi 2005		aktivoitunut		Yhteensä
		ei työllistynyt	työllistyi	
verrokit	ei työllistynyt	245	17	262
	työllistyi	30	3	33
Yhteensä		275	20	295

Vuoden 2004 työllistymisille eriäviä pareja, joissa vain toinen työllistyi, oli 39 kappaletta, joista 16:ssa aktivoitu työllistyi mutta verrokki ei. Testisuure on noudattaa χ^2 -jakaumaa, jolloin saatiin $p = 0,337$. Testin perusteella vertailtavat ryhmät eivät eronneet toisistaan työllistymisen suhteen. Vuoden 2005 tulokset olivat samansuuntaisia. Eriäviä pareja oli tuolloin 47 ja $p = 0,080$, joten ryhmät eivät eronneet toisistaan tilastollisesti.

Verrokit työllistyivät molempina vuosina paremmin kuin aktivoitunut, kuten alustavissa tarkasteluissakin. Tämän perusteella aktivointi näyttäisi jopa vähentävän työllistymistä, kun aktivointiin valikoituminen huomioitiin kaltaistumuuttujien osalta. Toisaalta on vielä selvitettävä tuloksen herkkyys havaitsemattomien tekijöiden vaikutukselle. Joka tapauksessa työllistyminen sinänsä on tutkimusjoukossa harvinainen tapahtuma: molempina vuosina yli 80 prosentissa pareista kumpikaan ei työllistynyt.

5.3.4 Sensitiivisyysanalyysi

Piilevän harhan vaikutusta McNemarin testin työllistymistuloksiin vuodelle 2004 selvitetiin sensitiivisyysanalyysillä. Jos oletetaan, ettei piilevää harhaa ole, McNemarin testi vertaa $T = 16$ työllistymistä aktivoitujen joukossa binomijakaumaan, jossa on 39 koetta ja onnistumisen todennäköisyys on $\frac{1}{2}$. Sensitiivisyysanalyysillä tutkittiin kuinka suuri vaikutus havaitsemattomilla tekijöillä on oltava, jotta tulokset eivät olisi merkitseviä. Tässä tapauksessa testauksessa saadut tulokset eivät kuitenkaan olleet alun perinkään merkitseviä, joten Rosenbaumin analyysi ei ollut suoraan sovellettavissa. Tulosta tarkasteltiin kahdella tapaa. Ensin tarkistettiin olisiko tulos ollut herkkä, jos se olisi ollut merkitsevä. Tämän jälkeen arvioitiin kuinka herkkä oli saatu tulos $T = 16$.

Jotta tulokset olisivat olleet lähes merkitseviä ($p = 0.053$), olisi pareja joissa vain aktivoitu työllistyi tullut olla vähintään $T = 26$. Jos näin olisi ollut, olisi valittu $e^g = \Gamma \geq 1$. Satunnaistetussa kokeessa $g = 0$, joten $p^+ = p^- = 1/2$, ja saadaan McNemarin testin merkitsevyytaso. Merkitsevyyden ylä- ja alarajojen funktioiden avulla laskettiin Γ arvot merkitsevyyksrajalla $\alpha = 0.050$. Tällöin ylärajan funktion ja merkitsevyyksrajan leikkauspisteeksi saatiin $\Gamma = 1.093$, joka on hyvin lähellä nollahypoteesitilanteen $\Gamma = 1$ arvoa. Koe olisi tällöin erittäin sensitiivinen, ja jo pienetkin poikkeamat havaitsemattoman tekijän suhteen vaikuttaisivat testin tulokseen. Tämän tarkastelun perusteella saatu tulos on hyvin epävarma, sillä työllistymiseen vaikuttavat useat seikat joista huomioimme vain rekisteristä löytyneet. Tiedämme siis, että on olemassa muuttujia jotka vaikuttavat työllistymiseen mutta joita meillä ei ollut mahdollisuutta käyttää, ja nämä saattaisivat helposti muuttaa testin tuloksen merkitseväksi.

Tämän jälkeen tarkasteltiin aineistosta saadun tuloksen sensitiivisyyttä. Koska havaittu $T = 16$, ei aktivoinnin vaikutusta voitu selvittää Rosenbaumin sensitiivisyysanalyysin mukaisesti, mutta sitä apuna käyttäen arvioimme tuloksen herkkyyttä käänteisessä tilanteessa. Tämä tarkoittaa, että jos saatu testin tulos on herkkä Rosenbaumin analyysin mukaan, voisivat mahdolliset työllisyyttä estävät seikat vaikuttaa siihen.

Tällöin T^+ ja T^- noudattivat binomijakaumaa 39 kokeella, joissa onnistumisen todennäköisyydet olivat

$$p^+ = \frac{\Gamma}{1+\Gamma} \text{ ja } p^- = \frac{1}{1+\Gamma}.$$

Kun nollahypoteesina on, ettei aktivointi vaikuta työllistymiseen, jokaiselle $g \geq 0$ saatiin

$$P(T^+ \geq a) \geq P(T \geq a | m) \geq P(T^- \geq a)$$

$$\Rightarrow \sum_{a=16}^{39} \binom{39}{a} (p^+)^a (1-p^+)^{39-a} \geq P(T \geq 16 | m) \geq \sum_{a=16}^{39} \binom{39}{a} (p^-)^a (1-p^-)^{39-a}$$

Kun ala- ja ylärajojen funktiot taulukoidaan eri Γ arvoilla, saadaan arvio sensitiivisyydestä Rosenbaumin tapaan. Taulukossa 11 näkyvät saadut sensitiivisyysanalyysin rajat. Lähtökohtana on nollahypoteesitilanne $\Gamma = 1$, joka on rajojen leikkauspiste. Alaraja leikkaa merkitsevän rajan kun $2.5 \leq \Gamma \leq 3.0$. Tämän jälkeen merkitsevä tulos on mahdollinen. Kun $\Gamma = 3$, toinen vastinparista saattaa olla aktivoitu jopa kolme kertaa todennäköisemmin kuin toinen koska heillä on erilaiset arvot havaitsemattoman kovariaatin u suhteen. Jotta kaksisuuntainen vastahypoteesi olisi uskottava, tulisi löytyä havaitsematon tekijä joka vaikuttaisi aktivoinnin ja työllistymisen vaikutussuhteeseen jonkin verran. Tämän perusteella tulos ei olisi niin sensitiivinen kuin edellisellä tapaa arvioituna, mutta on silti mahdollista että löytyy sellaisia havaitsemattomia tekijöitä jotka muuttaisivat kokeen tuloksen.

Taulukko 11. Sensitiivisyysanalyysin ylä- ja alarajat.

Γ	alaraja	yläraja
1.0	0.900	0.900
1.5	0.509	0.995
2.0	0.196	1.000
2.5	0.065	1.000
3.0	0.020	1.000
3.5	0.007	1.000
4.0	0.002	1.000

5.4 Yhteenveto ja tulosten tulkinta

Vaikuttavuuden arvioinnissa on oleellista selvittää sekä intervention vaikutuksen tulokset että prosessit (Dahler-Larsen ja Krogstrup 2003). Tutkielmassa pyrittiin selvittämään missä määrin saavutetut tulokset työllistymisessä johtuvat interventiosta, eli aktivointisuunnitelmien laadinnasta. Aktivoinnin prosesseista ja tuloksista on Stakes julkaissut aiemmin AKKU -tutkimuksen (Ala-Kauhaluoma ym. 2004).

Sovellusaineisto muodostui aktivointiin oikeutetuista pitkä-aikaistyöttömistä. Perusvuoden 2003 kuluessa aktivoidut muodostivat koeryhmän, kun kontrolliryhmässä yksilöille ei laadittu aktivointisuunnitelmia lainkaan. Työllistymistä seurattiin kahden seuraavan vuoden aikana. Henkilön katsottiin työllistyneen, jos hän oli työssä avoimilla työmarkkinoilla vähintään 90 päivää vuodessa.

Vaikuttavuutta tutkittiin tilastotieteellisesti kausaalisuuden näkökulmasta. Tähän pyrittiin kaltaistamalla vertailuryhmät propensiteettipistemäärän avulla. Sekoittavat kovariaatit saatiin näin tasapainotettua vertailuryhmissä, joten kaltaistus lähimmän naapurin menetelmällä oli onnistunut. Tosin näin vain havaitut tekijät tasapainoutuivat.

Aktivointitodennäköisyydet olivat hyvin pieniä. Myös työllistyminen oli harvinainen tapahtuma tutkimusjoukossa. Ryhmien väliset työllistymiserot olivat kaltaistuksen jälkeen hyvin lähellä alkuperäisestä aineistosta laskettuja. Tämä ei tarkoita että kaltaistus olisi ollut tarpeeton, vaan se osoittaa eron todelliseksi. Koeryhmä, eli vuonna 2003 aktivoidut, olivat prosenttiosuuksia vertailemalla huonommin työllistyviä kuin kaltaistettu verrokkiryhmä. Koeryhmästä työllistyi vuonna 2004 6,8 %, kun kaltaistetusta kontrolliryhmästä työllistyi 9,2 %. Alkuperäisestä aineistosta suoraan laskettu supistetun kontrolliryhmän (yhteensä 14054 henkilöä) työllistyneiden osuus oli 16,1 %.

Käsittelyvaikutusta tutkittiin McNemarin riippuvien ryhmien parametrittomalla testillä. Kaltaistettujen ryhmien työllistymisessä ei sen perusteella ollut eroa kun vastahypoteesi oli kaksisuuntainen. Ilmeisesti aktivoitaviksi oli valikoitu yleisesti ottaen erityisen huonosti työllistyviä henkilöitä, eikä heidän työllistymistään saatu parannettua aktivointisuunnitelmien laadinnalla ei-

valikoituja paremmaksi. Voimakasta valikoitumista on havaittu aikaisemmissakin tutkimuksissa (esim. Tupo 2 –työryhmän mietintö 2006).

Sensitiivisyysanalyysillä kartoitettiin mahdollisen piilevän harhan vaikutusta tuloksiin. Koska testitulos ei ollut alun perinkään merkitsevä, ei Rosenbaumin sensitiivisyysanalyysiä voitu soveltaa suoraan, ja sensitiivisyyttä tarkasteltiin kahdella tapaa Rosenbaumin analyysiä soveltaen. Kun parien määrä joissa ainoastaan aktivoitu työllistyi, nostettiin 26:een, saatiin keinotekoinen tilanne jolloin tulos olisi ollut merkitsevä. Analyysin perusteella saatu tulos oli erittäin sensitiivinen, joten se olisi herkkä piilevien tekijöiden vaikutukselle. Havaittua tulosta $T = 16$ analysoitiin vielä erikseen, jolloin osoittautui, että havaitsemattomien tekijöiden tulisi olla vahvoja työllistymistä estäviä tekijöitä, jotta ne muuttaisivat tulokset merkitseviksi. Etukäteen oli odotettavaa että havaitsemattomia tekijöitä olisi, ja niiden vaikutus saattaisi olla huomattavaakin.

Tutkielmassa selvitettiin havainnoivan tutkimuksen harhanlähteitä ja niiden kontrollointimahdollisuuksia. Yhteiskuntatieteellisissä asetelmissa tutkittavan asian kokonaisvaltaisuus on tilastotieteellisessä mielessä ongelma. Aikaisempien tutkimusten perusteella aktivointisuunnitelmien laadinnalla on hyvinvointia lisääviä vaikutuksia. Kuntouttavan työtoiminnan ohjausryhmän muistiossa todettiin, että lain toimenpiteet vahvistivat viranomaisyhteistyötä, ja että kuntouttavasta työtoiminnasta tuli väline työttömien elämänhallinnan parantamiseksi. (mm. Kuntouttavan työtoiminnan ohjausryhmä 2004: 10.) Sovelluksessa saadun uuden tiedon perusteella aktivointisuunnitelmilla ei ollut positiivista vaikutuksista pitkäaikaistyöttömien työllistymiseen seurantavälillä. Aikaisemmin on todettu, että tavallisesti toimenpiteiden arviointiin käytetyt ei –kokeelliset menetelmät yliarvioivat niiden vaikutuksen työllistymiseen (Hämäläinen ym. 2007). Nyt saatu tulos osoitti, että aktivointisuunnitelmien osalta tulisi löytyä tässä havaitsematta jääneitä tekijöitä jotta tulokset olisivat merkitseviä. Tulos onkin herkkä havaitsemattomien työllistymistä estävien tekijöiden vaikutuksille. Tulokset sopivat yhteen aikaisempien tutkimusten kanssa, vaikka kaltaistusmenetelmiä on Suomessa käytetty sosiaalipolitiikan saralla vielä vähän.

6. Johtopäätökset

Tutkielman tavoite oli selvittää aktivointisuunnitelmien laadinnan vaikutuksia työllistymiseen. Satunnaistetussa koetilanteessa myös kausaalisuhteeseen vaikuttavat piilevät tekijät tasapainottuisivat, joten se olisi harhattomin tapa tarkastella työllistymistä, mutta tämän tutkimuskysymyksen puitteissa käytännössä mahdoton. Havainnoivissa asetelmissa on luotava tilanne, joka vastaa yksilöiden tuntemattomien potentiaalisten vasteiden tutkimista. Tällöin on huomioitava käsittelyyn valikoituminen, joka aktivointisuunnitelmien tapauksessa oli merkittävää. Käsittelyryhmät kaltaistettiin valikoitumisen perusteella. Kaltaistuksessa käytetyn ennustetun todennäköisyyden, eli propensiteettipistemäärän, ja lähimmän naapurin menetelmän avulla saatiin muodostettua havaittujen tekijöiden suhteen vertailukelpoiset ryhmät. Kaltaistuksen jälkeen vertailuryhmät eivät eronneet työllistymisen suhteen, joten lain työllistämistavoitteet eivät tulosten perusteella toteutuneet. Aineistolla olisi mahdollista tutkia tulosta myös muiden kaltaistusmenetelmien avulla.

Koska koe oli kuitenkin sensitiivinen, olisi aineistoon hyvä saada mukaan ainakin sosiaalisia ongelmia ja mahdollisesti epävirallisten suhteiden puuttumista kuvaavia muuttujia. Tilastollisesti on mahdollista vain osoittaa havaitsemattomien tekijöiden mahdollinen vaikutus, mutta ei arvioida onko niitä todella tai mitä ne voisivat olla. Kokeen sensitiivisyydestä johtuen tuloksia on tulkittava varauksella. Olisi mielenkiintoista tutkia asetelmaa myös etenevänä, joka ei nyt ollut mahdollista työllistymisen määrittelystä johtuen.

Lähdeluettelo

Ala-Kauhaluoma, Mika, Keskitalo, Elsa, Lindqvist, Tuija ja Parpo, Antti (2004): *Työttömien aktivointi*. Kuntouttava työtoiminta –lain sisältö ja vaikuttavuus. Stakes, tutkimuksia 141. Helsinki.

Aho, Simo (2004): *Kroonisen työttömyyden laajuus, rakenne ja syntytausta*. Työministeriö, työpoliittinen tutkimus 261. Helsinki.

Aho, Simo ja Koponen, Hannu (2007): *Työvoimapolitiittisiin toimenpiteisiin osallistuneiden seuranta 2*. Tutkimus rekisteriaineisosta vuosilta 2002-2005. Työministeriö, työpoliittinen tutkimus 324. Helsinki.

Aho, Simo ja Kunttu, Susanna (2001): *Työvoimapolitiittisten toimien vaikuttavuuden tutkiminen rekisteriaineistojen avulla*. Menetelmäkehittelyä ja 1994-1997 toteutettuja toimia koskevia tuloksia. Työministeriö, työpoliittinen tutkimus 233. Helsinki.

Beagerhole, R., Bonita, R. ja Kjellström, T. (1993): *Basic Epidemiology*. World Health Organization. Geneve.

Clayton, David ja Hills, Michael (1993): *Statistical Models in Epidemiology*. Oxford University Press. Iso-Britannia.

Cox, David R. (1972): Regression models and life tables. *Journal of the Royal Statistical Society*. Series B 34:187-220. London.

Deheija Rajeev H. ja Wahba Sadek (2002): Propensity Score Methods for Nonexperimental Causal Studies. *The review of Economics and Statistics* 84(1), 151-161.

Dahler-Larsen, Peter ja Krogstrup, Hanne K. (2003): *Nye Veje i Evaluering*. Århas. Systime. Osakäännös: FinSoc Arviointiraportteja (3/2005): Vaikuttavuuden arviointi. Stakes. Helsinki.

Fisher, R.A. (1935) *The Design of Experiments*. Oliver & Boyd. Edinburgh.

Ho, Daniel E., Imai, Kosuke, King, Gary ja Stuart, Elizabeth A. (2007): Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Political Analysis*: 15, 199-236. Saatavilla www-muodossa: <URL: <http://gking.harvard.edu/files/matchp.pdf>>

Hämäläinen, Kari (1998): *Aktiivinen työvoimapolitiikka Suomen työmarkkinoilla*. Taloustieteen näkökulmia ja tutkimustuloksia. Työministeriö, työpoliittinen tutkimus 192. Helsinki.

Hämäläinen, Kari, Nykyri, Elina, Uusitalo, Roope ja Vuori, Jukka (2007): *Ovatko vaikuttavuusarviot kohdallaan*. Esimerkkinä työnhakukoulutus. Työministeriö, työpoliittinen tutkimus 329. Helsinki.

Hämäläinen, Kari ja Tuomala, Juha (2006): *Työvoimapolitiittisten toimenpiteiden vaikutusten arviointi*. Työministeriö, työpoliittinen tutkimus 315. Helsinki.

Kuntouttavan työtoiminnan ohjausryhmä (2004): *Kuntouttavan työtoiminnan ohjausryhmän muistio*. Sosiaali- ja terveysministeriön työryhmämuistioita 2004:10. Sosiaali- ja terveysministeriö. Helsinki.

Lennon, Eeva (2005): Köyhyys ja työttömyys alaspäin Britanniassa. *Yhteiskuntapolitiikka*: 70, 4/2005, 466-468.

Luukkonen, Antti, Korkeamäki, Ossi ja Kyyrä, Tomi (2005): *Aktiivinen työvoimapolitiikka yritysten näkökulmasta*. Työministeriö, työpoliittinen tutkimus 270. Helsinki.

Paul W. Holland (1986): Statistics and Causal Inference. *Journal of American Statistical Association*: 81, 945-960.

Pearl, Judea (1997): Graphical Models for Probabilistic and Causal Reasoning. Teoksessa Tucker, Allen B. (toim.) (1997, 697-714): *The Computer Science and Engineering Handbook 1997*. CRC Press.

Peirce, Charles Sanders (1868): Perusteita logiikan lakien pätevyydelle. Teoksessa Lång, Markus (suom.) (2001, 90-124): *Johdatus tieteen logiikkaan ja muita kirjoituksia*. Vastapaino. Tampere.

Rosenbaum, Paul R. (1995): *Observational Studies*. Springer-Verlag. New York.

Rosenbaum, Paul ja Rubin, Donald B. (1983): The central role of the propensity score in observational studies for causal effects. *Biometrika*: 70: 41-55.

Rubin, Donald B. (1977): Assignment to Treatment Group on the Basis of a Covariate. *Journal of Educational Statistics*: Vol. 2, No. 1, 1-26.

Rubin, Donald B. (1997): Estimating Causal Effects from Large Data Sets Using Propensity Scores. *Annals of Internal Medicine*: Vol. 127, 757-763.

Seidenfeld, T. (1979): *Philosophical problems of statistical inference*. Learning from R. A. Fisher. D. Reidel Publishing Company. Dordrecht.

Soininvaara, Osmo (1999): *Täystyöllisyyteen ilman köyhyyttä*. Art House. Helsinki.

Työllisyysohjelman loppuraportti (3/2007): *Työllisyys nousussa*. Hallituksen politiikkaohjelmat, työllisyys.

TE -keskukset [Viitattu 1.8.2006]. Saatavilla www -muodossa: <URL: <http://www.te-keskus.fi/web/telap.nsf/FrameSetFIN?OpenFrameSet>>

Tuominen, Pekka (1993): *Todennäköisyyslaskenta*. Limes. Helsinki.

Tupo 2 - työryhmän mietintö (2006): *Väyliä työhön*. Työministeriö, työhallinnon julkaisu 361. Helsinki.

Työministeriö [viitattu 18.8.2007]. Saatavilla www -muodossa <URL: http://www.mol.fi/mol/fi/06_tyoministerio/07_tilastot/02_aikasarjat/01_tyonhakijat/index.jsp>

Liitteet

LIITE 1. Riski ja uskottavuuspäätely

Todennäköisyydellä p kuvataan riskiä tapahtuman toteutumiselle. Esimerkiksi binäärisen vasteen tapahtuman todennäköisyys on $p = P(Y = 1)$. Merkitään tapahtuman välttötodennäköisyyttä $(1 - p)$:llä. Kun tarkastellaan riskejä tutkimusasetelman koe- ja kontrolliryhmissä, ovat ne ehdollisia käsittelylle T

$$p_T = P(Y = 1|T = 1) \text{ ja } p_C = P(Y = 1|T = 0).$$

Riskiä voidaan estimoida tapausten määrällä suhteessa kaikkiin mahdollisuuksiin. Tähän päästään uskottavuuspäätelyn kautta.

Riskin malli N :lle yksilölle voidaan esittää Bernoulli uskottavuusfunktiona. Mikäli D_T tapahtumaa toteutuu koeryhmässä, on riskin p_T uskottavuusfunktio muotoa

$$L(p_T) = (p_T)^{D_T} (1 - p_T)^{N_T - D_T},$$

jossa N_T on koeyksilöiden kokonaismäärä. Kontrolliryhmälle on vastaava uskottavuusfunktio $L(p_C)$. Log-uskottavuusfunktio koeryhmälle on muotoa

$$\log L(p_T) = D_T \log(p_T) + (N_T - D_T) \log(1 - p_T).$$

Ratkaistaan uskottavuusyhtälö p_T :n suhteen

$$\frac{\partial}{\partial p_T} \log L(p_T) = \frac{D_T}{p_T} + \frac{(N_T - D_T)}{1 - p_T} = 0,$$

jolloin suurimman uskottavuuden estimaattoriksi saadaan

$$\hat{p}_T = \frac{D_T}{N_T}.$$

Riskiä p_T voidaan siis arvioida tapausten ja kaikkien käsittelyn saaneiden yksilöiden lukumäärän suhteella. (Clayton & Hills 1993: 43-45)

LIITE 2: Muuttajat ja työllistymisen aktivoinnin suhteen vuonna 2004

muuttaja	luokitus	ei -aktivoidut n= 42890		koeryhmä n=295		Yhteensä
		ei työssä 2004	työssä 2004	ei työssä 2004	työssä 2004	
sukupuoli	mies	17537 (84,4%)	3250 (15,6%)	162 (91,5%)	15 (8,5%)	20964
	nainen	18359 (83,1%)	3744 (16,9%)	113 (95,8%)	5 (4,2%)	22221
työnhakuala	toimisto	15495 (82,7%)	3245 (17,3%)	81 (93,1%)	6 (6,9%)	18827
	teollisuus	9435 (80,1%)	2340 (19,9%)	105 (92,9%)	8 (7,1%)	11888
	palvelu/muu	10966 (88,6%)	1409 (11,4%)	89 (93,7%)	6 (6,3%)	12470
työssä 2002	työllistyi	32416 (88,1%)	4383 (11,9%)	266 (94,0%)	17 (6,0%)	37082
	ei työllistynyt	3480 (57,1%)	2611 (42,9%)	9 (75,0%)	3 (25%)	6103
vajaakuntoisuus	ei	32667 (83,3%)	6544 (16,7%)	187 (91,7%)	17 (8,3%)	39415
	on	3229 (87,8%)	450 (12,2%)	88 (96,7%)	3 (3,3%)	3770
sijoitus	ei	33400 (83,8%)	6478 (16,2%)	226 (93,8%)	15 (6,2%)	40119
	on	2496 (82,9%)	516 (17,1%)	49 (90,7%)	5 (9,3%)	3066
kurssi	ei	34187 (84,0%)	6536 (16,0%)	245 (93,2%)	18 (6,8%)	40986
	on	1709 (78,9%)	458 (21,1%)	30 (93,8%)	2 (6,3%)	2199
koulutus	perus	11899 (87,6%)	1684 (12,4%)	151 (96,8%)	5 (3,2%)	13739
	keskiaste	15383 (81,6%)	3469 (18,4%)	97 (89,0%)	12 (11,0%)	18961
	korkea	8614 (82,4%)	1841 (17,6%)	27 (90,0%)	3 (10,0%)	10485
Yhteensä		35896(83,7%)	6994 (16,3%)	275 (93,2%)	20 (6,8%)	43185

*taulukossa ilmoitettu havaintojen kulumäärä, sulkeissa prosentiosuus käsittelyryhmän ja muuttujan luokan suhteen.

muuttaja	luokitus	ei -aktivoidut n= 40059		koeryhmä n=295		Yhteensä
		ei työssä 2004	työssä 2004	ei työssä 2004	työssä 2004	
ikäluokka	18-24	7364 (86,3%)	1167 (13,7%)	52 (96,3%)	2 (3,7%)	8585
	25-34	7690 (81,9%)	1704 (18,1%)	37 (78,7%)	10 (21,3%)	9441
	35-44	6292 (77,8%)	1791 (22,2%)	66 (91,7%)	6 (8,3%)	8155
	45-54	5637 (77,8%)	1601 (22,2%)	80 (97,6%)	2 (2,4%)	7320
	55-	6271 (92,0%)	542 (8,0%)	40 (14,5%)	0 (0%)	6853
Yhteensä		33254 (83,0%)	6805 (17,0 %)	275 (93,2%)	20 (6,8%)	40354

Huom. Ikäluokasta osa on tässä jätetty huomiotta, sillä he eivät mahdu valittuun luokitukseen, joten kokonaissummat ovat pienempiä.

LIITE 3: Ennustettujen todennäköisyyksien jakauma

Ennustettujen todennäköisyyksien jakauma

