

Elias Einiö ja Ari Hyytinen

Yritystukien vaikuttavuuden arviointi satunnaistettujen vertailukokeiden avulla

Valtioneuvoston
selvitys- ja tutkimus-
toiminnan julkaisusarja

2019:45

ISSN 2342-6799

ISBN PDF 978-952-287-765-9

Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja 2019:45

Yritystukien vaikuttavuuden arviointi satunnaistettujen vertailukokeiden avulla

Elias Einiö ja Ari Hyytinen

Valtioneuvoston kanslia

ISBN PDF: 978-952-287-765-9

Helsinki 2019

Kuvailulehti

Julkaisija	Valtioneuvoston kanslia	27.6.2019
Tekijät	Elias Einiö ja Ari Hyytinen	
Julkaisun nimi	Yritystukien vaikuttavuuden arviointi satunnaistettujen vertailukokeiden avulla	
Julkaisusarjan nimi ja numero	Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja 2019:45	
ISBN PDF	978-952-287-765-9	ISSN PDF 2342-6799
URN-osoite	http://urn.fi/URN:ISBN:978-952-287-765-9	
Sivumäärä	53	Kieli suomi
Asiasanat	Yritystuet, satunnaistetut vertailukokeet, satunnaistaminen, tutkimus, tutkimustoiminta	
Tiivistelmä:	<p>Toisin kuin monilla muilla yhteiskunnan alueilla (mm. lääketiede ja sosiaaliturva) yritystukien vaikutuksia ei ole Suomessa tutkittu kontrolloiduilla ja etukäteen suunnitelluilla tieteellisillä kokeilla. Tämä siitakin huolimatta, että yritystukien kohdalla satunnaistettujen kontrolloitujen arviointikokeiden (randomized controlled trial, RCT) toteuttamiselle ei välttämättä ole suuria lainsäädännöllisiä tai teknisiä esteitä. Tuen vaikuttavuudella tarkoitetaan tässä nimenomaan aitoa syy-seuraussuhdetta, kuten esimerkiksi tuen kausaalivaikutusta yrityksen T&K-toimintaan tai sen työllistämien ihmisten määrään, eikä esimerkiksi sitä, kuinka hyödylliseksi tukea saaneet yritykset ovat tuen itse (subjektiivisesti) kokeneet. Tässä raportissa esittelemme satunnaistettuja kontrolloituja kokeita yritystukien vaikuttavuusarviointiin näkökulmasta. Satunnaistetut koeasetelmat ovat luotettava ja siksi suositeltava arviointimenetelmä suurelle osalle yritystuki-instrumenteista. Kokeiden käytännön järjestelyihin liittyy kuitenkin lukuisia yksityiskohtia, jotka riippuvat ohjelmien tavoitteista, institutionaalisista erityispiirteistä ja tarvittavista aineistoista. Kokeiden suunnittelu edellyttää siksi aina tiivistä yhteistyötä tutkijoiden ja ohjelmantoteuttavien tahojen välillä. Yritystukiohjelmien vaikuttavuusarviointiin koeasetelmien avulla liittyy myös erityiskysymyksiä. Kuten tässä raportissa kuvataan, niitä ovat mm. 1) ulkoisvaikutusten mittaaminen ja huomioonotto, 2) tuen saajan omien toimien merkitys vaikuttavuusmittaukselle ja 3) erilaisten vaikutusmekanismien tunnistamiseen liittyvät haasteet. Myös vaikuttavuusarvioiden mahdollinen kontekstisidonnaisuus ja tuki-instrumentteihin liittyvät skaalautuvuus-kysymykset on syytä ottaa huomioon arvioita tehtäessä ja tulkittaessa. Esittelemme systemaattisen yritystukien jaottelukehikon, jonka avulla tuet voidaan luokitella niihin, jotka voidaan arvioida satunnaistetuilla koeasetelmilla ja niihin, jotka täytyy arvioida muulla tavoin. Ehdotuksemme on, että tähän luokitteluun pohjautuen tietyjen yritystukiohjelmien vaikutuksia tulisi jatkossa arvioida aika-ajoin kontrolloiduilla satunnaistetuilla koeasetelmilla. Tämän ohella ehdotamme systemaattisen arvioinnin käynnistämistä kahdella pilottikokeella. Toinen piloteista koskee Business Finlandin T&K-tukia ja toinen ELY-toimistojen jakamaa työllistämistukea. Pohdimme lyhyesti myös sellaisten tukien arviointia, joihin tavanomainen satunnaistettu koeasetelma soveltuu huonommin, kuten niiden, joissa tukea saa vain muutama suuri yritys. Kirjoituksen tarkoituksena on antaa lukijalle kattava mutta tiivis kokonaiskuva yritystukien satunnaistettuihin arviointitutkimuksiin liittyvistä tärkeistä käsitteistä ja kokeen onnistumisen kannalta keskeisistä kysymyksistä. Yksi tämän raportin tavoitteista on, että se osaltaan lisää keskustelua, jota yritystukien vaikuttavuusarvioinneista päättävät, tukiohjelmiä toteuttavat tahot ja tutkijayhteisö aiheesta käyvät.</p>	
Tämä julkaisu on toteutettu osana valtioneuvoston selvitys- ja tutkimussuunnitelman toimeenpanoa. (tietokayttoon.fi) Julkaisun sisällöstä vastaavat tiedon tuottajat, eikä tekstisisältö välttämättä edusta valtioneuvoston näkemystä.		
Kustantaja	Valtioneuvoston kanslia	
Julkaisun myynti/jakaja	Sähköinen versio: julkaisut.valtioneuvosto.fi Julkaisumyynti: julkaisutilaukset.valtioneuvosto.fi	

Presentationsblad

Utgivare	Statsrådets kansli	27.6.2019	
Författare	Elias Einiö och Ari Hyytinen		
Publikationens titel	Evaluering av företagsstöds påverkan med hjälp av randomiserade och kontrollerade studier		
Publikationsseriens namn och nummer	Statsrådets publikationsserie för utrednings- och forskningsverksamheten 2019:45		
ISBN PDF	978-952-287-765-9	ISSN PDF	2342-6799
URN-adress	http://urn.fi/URN:ISBN:978-952-287-765-9		
Sidantal	53	Språk	finska
Nyckelord	Företagsstöd, randomiserade och kontrollerade studier, forskning, forskningsverksamhet		
Referat:	<p>Till skillnad från många andra områden i vårt samhälle (t. ex. medicin och socialskydd) har man inte utrett företagsstöds påverkan med kontrollerade och vetenskapliga studier som har planerats på förhand. Det är så även om det inte finns särskilda juridiska eller tekniska hinder för att randomiserade och kontrollerade studier (randomized controlled trial, RCT) kunde användas när företagsstöd evalueras. Företagsstöds påverkan betyder i detta fall de äkta orsakssamband som ligger bakom stödets effekt på företags utvecklings- och forskningssatsningar eller på rekryteringar. Företagsstöds påverkan hänvisar inte till hur de företag som har fått stöd subjektivt upplever att stödet har gynnat dem. I denna rapport beskriver vi hur randomiserade och kontrollerade studier kunde användas när företagsstöds påverkan utreds. Randomiserade och kontrollerade studier är pålitliga och de är därför en tillräddlig evalueringsmetod för många olika stödformer. I praktiken finns det många detaljer som man måste ta hand om innan randomiserade och kontrollerade studier kan utnyttjas. Detaljerna beror på syftet med stödprogram, särdrag hos relevanta institutioner och tillgängliga datamaterial. Planering av studier förutsätter därför samarbete mellan forskare och de som är ansvariga för stödprogrammen. Om man vill använda kontrollerade studier för att evaluera stödprogrammets påverkan måste man också ta ställning till några särskilda frågor. Såsom vi beskriver i denna rapport inkluderar de här frågorna bl. a. 1) hur man kan mäta och ta hänsyn till yttre effekter, 2) hur stötagares beteende påverkar mätning, och 3) hur man kunde identifiera mekanismer som ligger bakom effekter. När man genomför och tolkar evalueringarna måste man också beakta att stödets påverkan kan bero på miljön och att det kan vara utmanande att implementera randomiserade evalueringar av stödprogram i större skala. Vi introducerar en systematisk analysram för företagsstöd som kan användas för att dela olika stödformer till dem som borde evalueras med randomiserade kontrollerade studier och till dem som måste evalueras på ett annat sätt. Vi föreslår att ramen används för identifiering av de stödprogram som i fortsättningen borde evalueras systematiskt med randomiserade kontrollerade studier. Vi föreslår också att den systematiska evalueringen lanseras i två pilotprojekt. I det första pilotprojektet skulle Business Finlands utvecklings- och forskningsstöd evalueras och i det andra pilotprojektet skulle påverkan av sysselsättningsstöden som ELY centraler allokera evalueras. Syftet med den här rapporten är att ge en omfattande men kompakt beskrivning av de viktiga begrepp som ansluter till randomiserade och kontrollerade evalueringsstudier och av de frågor som är avgörande för att sådana studier kan genomföras framgångsrikt. Ett särskilt syfte med rapporten är att främja debatt mellan de som bestämmer om evalueringar, de som är ansvariga för stödprogrammen och de som forskar stödprogrammets påverkan.</p>		
	Den här publikation är en del i genomförandet av statsrådets utrednings- och forskningsplan. (tietokaytoon.fi) De som producerar informationen ansvarar för innehållet i publikationen. Textinnehållet återspeglar inte nödvändigtvis statsrådets ståndpunkt		
Förläggare	Statsrådets kansli		
Beställningar/ distribution	Elektronisk version: julkaisut.valtioneuvosto.fi Beställningar: julkaisutilaukset.valtioneuvosto.fi		

Description sheet

Published by	Prime Minister's Office	27 June 2019	
Authors	Elias Einiö and Ari Hyytinen		
Title of publication	Use of randomized controlled trials in the evaluation of the impact of business subsidies		
Series and publication number	Publication series of the Government's analysis, assessment and research 2019:45		
ISBN PDF	978-952-287-765-9	ISSN PDF	2342-6799
Website address URN	http://urn.fi/URN:ISBN:978-952-287-765-9		
Pages	53	Language	Finnish
Keywords	Business subsidies, randomized controlled trials, research, research activities		
<p>Abstract:</p> <p>Unlike in many other areas of society (e.g., medicine and social insurance), the effects of company support have not been previously studied with controlled and pre-planned scientific trials in Finland. This is the case, even though there are no large legal or technical restrictions prohibiting the use of randomized control trials. In this report, the effectiveness of company support means a true causal relationship, for example, the causal impact of the support on the company's R&D activity or on the number of people it employs. The report presents randomized control trials from the perspective of company support impact evaluation. Randomized controlled trials are reliable, and therefore, they can be recommended as an evaluation method for most company support instruments. The practical execution of the trials involves many details, which depend on the program goals, institutional characteristics and the required data. The planning of the trials always requires close collaboration between researchers and parties executing the program. There are a number of special issues relating to the evaluation of company support using trials. As described in this report, these include 1) accounting for and measuring externalities, 2) the effect of the support recipients' actions on the evaluation inference and 3) challenges in identifying different impact mechanisms. Also, the possible context-dependency and scaling issues with support instruments should be taken into account when executing and interpreting the evaluations. We present a systematic framework which can be used to classify the company support instruments into those which can be assessed using randomized trials and those which must be evaluated using other methods. Based on this categorization, our proposal is that in the future the effects of certain company support instruments should be evaluated from time to time with randomized control trials. In addition to this, we propose the commencement of systematic evaluation with two pilot studies. The first pilot examines Business Finland's R&D subsidies and the other employment subsidies awarded by the Centre for Economic Development, Transport and the Environment. We also briefly consider the evaluation of support, which cannot be readily examined using regular randomized control trials, such as subsidies that are only received by a few large companies. The purpose of this report is to provide the reader with a comprehensive, but compact treatment of important concepts related to randomized evaluations and questions that are essential for running a successful trial. One of the goals of this report is to generate dialogue between those who decide about company support, those who execute the programs and the research community.</p>			
<p>This publication is part of the implementation of the Government Plan for Analysis, Assessment and Research. (tietokaytoon.fi) The content is the responsibility of the producers of the information and does not necessarily represent the view of the Government.</p>			
Publisher	Prime Minister's Office		
Publication sales/ Distributed by	Online version: julkaisut.valtioneuvosto.fi Publication sales: julkaisutilaukset.valtioneuvosto.fi		

Sisältö

1	Johdanto	8
2	Yleistä satunnaistetuista kontrolloiduista koejärjestelyistä	12
2.1	Koe- ja verrokkiryhmä	12
2.2	Kokeen koko	13
2.3	Tutkimuskysymyksen ja -hypoteesien asettaminen	15
2.4	Kohdepopulaation valinta	18
2.5	Satunnaistaminen käytännössä	19
2.6	Satunnaistettava toimenpide	21
2.7	Vastemuuttajat: Rekisteriaineistot vs. tietojen kerääminen kyselyillä.....	22
2.8	Heterogeeniset vaikutukset ja tulosten yleistettävyys	23
2.9	Arvioinnin avoimuus ja toistettavuus	23
2.10	Arviointisuunnitelman etukäteisrekisteröinti	24
3	Erityiskysymyksiä	25
3.1	Erityiskysymys 1: Ulkoisvaikutukset.....	25
3.1.1	Miksi ulkoisvaikutuksiin pitää kiinnittää huomiota?	25
3.1.2	Ulkoisvaikutusten mittaamisen haasteita.....	27
3.2	Erityiskysymys 2: Koejärjestelyn toteuttaminen, kun toimenpiteen kohteeksi joutuvan omilla toimilla on merkitystä sille, miten toimenpide vaikuttaa	32
3.3	Erityiskysymys 3: Syyseuraus-suhteiden mekanismien ymmärrys	35
3.4	Muita erityiskysymyksiä.....	39
4	Satunnaistettu vaikuttavuusarviointi yritystukipolitiikan kehittämistyökaluna	41
4.1	Vaikuttavuusarvioinnin kehitys-arviointi -sykli	41
4.2	Käytännön näkökohtia tuki-instituutioiden osallistamisesta.....	42
4.3	Muut ekonometriset arviointimenetelmät.....	43

5	Johtopäätökset ja loppupohdintaa	46
5.1	Minkälaisia yritystukia olisi perusteltua arvioida satunnaistettujen koejärjestelyiden avulla?	46
5.2	Kohti järjestelmällistä vaikuttavuusarviointia?	48
	Lähteet	52

1 Johdanto

Suomessa maksetaan yritystukia vuosittain noin 4 mrd. euroa (TEM, 2017). Jos mukaan ei lueta esimerkiksi verotukia, vaan ainoastaan suorat tuet, lainat ja takaukset, yritystukea sai vuonna 2017 yhteensä reilut 28000 yritystä (Tilastokeskus 2018). Tässä raportissa tarkastellaan yritystukien arviointia kontrolloiduilla satunnaistetuilla koeasetelmilla (*randomized controlled trials, RCTs*). Tarkoitamme yritystukien arvioinnilla sekä verotukien että erilaisten suorien ja epäsuorien rahamääräisten tuki-instrumenttien, joihin sisältyvä tukiosuus maksetaan julkisista varoista, vaikutusten selvittämistä.

Toisin kuin monilla muilla yhteiskunnan alueilla (mm. lääketiede ja sosiaaliturva) yritystukien vaikutuksia ei ole Suomessa tutkittu kontrolloiduilla ja etukäteen suunnitelluilla tieteellisillä kokeilla. Tämä siitakin huolimatta, että yritystukien kohdalla satunnaistettujen kontrolloitujen arviointikokeiden (*randomized controlled trial, RCT*) toteuttamiselle ei välttämättä ole suuria lainsäädännöllisiä tai teknisiä esteitä. Tuen vaikuttavuudella tarkoitetaan tässä nimenomaan aitoa syy-seuraussuhdetta, kuten esimerkiksi tuen kausaalivaikutusta yrityksen T&K-toimintaan tai sen työllistämien ihmisten määrään, eikä esimerkiksi sitä, kuinka hyödylliseksi tukea saaneet yritykset ovat tuen itse (subjektiivisesti) kokeneet. Tuen vaikuttavuutta ei voida myöskään päätellä siitä, että tukea saavat yritykset pärjäävät jollakin valitulla mittarilla mitattuna paremmin kuin ne yritykset, jotka eivät saaneet tukea. On nimittäin täysin mahdollista, että yritystukia päätyvät hakemaan yritykset, jotka pärjäisivät hyvin myös siinä tapauksessa, että ne eivät saa tukea.

Satunnaistettujen koeasetelmien idea ei ole uusi. Lääketieteessä lääkkeiden vaikutuksia on pitkään tutkittu jakamalla kokeeseen osallistuvat henkilöt satunnaisesti koe- ja verrokkiryhmiin. Kun kokeessa on riittävän monta osallistujaa, tuottaa satunnaistaminen keskimäärin samanlaiset ryhmät, jotka ovat keskenään vertailukelpoisia. Verrokkiryhmän avulla voidaan päätellä, mitä koeryhmälle olisi tapahtunut ilman toimenpidettä eli ns. kontrafaktuaali. Tästä syystä koe- ja verrokkiryhmiä vertaamalla saadaan luotettavaa tietoa toimenpiteen vaikutuksesta.

Yritystukiohjelmien systemaattisella testaamisella satunnaistettujen arviointitutkimusten avulla on useita hyötyjä. Ilman niitä sellainen tukien kvantitatiivinen arviointi, jonka

tavoitteena on kausaalivaikutusten mittaaminen, jää usein näennäiskoeasetelmien (*quasi experiment*) varaan. Näennäiskoeasetelmilla tarkoitetaan tilannetta, jossa jonkin politiikkatoimen tai toimenpiteen piirre, kuten esimerkiksi tietyn yritystuki-instrumentin myöntöprosessi, tuottaa tahattomasti tilanteen, jossa sekä havaituilta ja havaitsemattomilta ominaisuuksiltaan samanlaisia yrityksiä saa tukea ja jää ilman tukea. Vaikka näennäiskoeasetelmat voivat tuottaa luotettavaa tietoa yritystukien vaikutuksista, ei niitä ole käytettävissä kaikkien yritystuki-instrumenttien kohdalla. Näin ollen näennäiskokeiden varassa oleva tietopohja ei välttämättä vastaa poliittisen päätöksenteon kannalta keskeisiin kysymyksiin tai kata riittävästi erilaisia tutki-instrumentteja. Vaihtoehtoisesti voidaan tutkia muutoksia tuettujen ja tukea ilman jäävien yritysten välillä, kontrolloiden yritysten välisiä havaittuja eroja, kuten mm. Laukkanen, Ollikka ja Tamminen (2019) ovat tehneet. Tällöin tuloksien luotettavuutta voi heikentää kuitenkin valikoituminen havaitsemattomien tekijöiden suhteen (esim. yritysten väliset erot panosten tai tuotteiden hintojen muutoksissa).

Satunnaistetut vaikuttavuusarviointit ovat vahvimmillaan suorien tai välittömien käytäytymisvaikutusten testaamisessa. Niillä tarkoitetaan muutosta tuettujen yritysten toiminnassa ja menestyksessä, verrattuna verrokkiryhmään. Suorien vaikutusten merkitys on ilmeinen. Olivat tukitoimenpiteen tavoitteet mitkä tahansa, niitä ei saavuteta, jos tukitoimella ei ole minkäänlaista vaikutusta tuettavien yritysten toimintaan.

Erityisesti yrityssektoriin kohdistetuilla toimilla on usein myös tavoitteita, joiden arviointiin ei riitä vaikutusten tarkastelu vain tuettujen yritysten osalta. Tämä seuraa siitä, että tuki voi vaikuttaa välillisesti, kuten mm. markkinamekanismien kautta, ilman tukea jääviin kilpailijoihin. Tällaisia ovat esimerkiksi tuote- ja tuotannontekijämarkkinoiden kautta syntyvät kilpailuvaikutukset ja innovaatioiden kohdalla teknologioita koskevan tiedon leviäminen. Tuella voi olla myös muunlaisia ulkoisvaikutuksia, jotka voivat olla positiivisia tai negatiivisia (ks. mm. Laukkanen ja Maliranta, 2019). Epäsuorien tai välillisten vaikutusten arviointi on yleisesti vaikeampaa, mutta uusien tutkimus osoittaa, että satunnaistetuilla tarkkaan rakennetuilla koeasetelmilla epäsuorien vaikutusten tarkastelu voidaan sen haastavuudesta riippumatta toteuttaa luotettavasti.

Yritystuki-instrumenttien satunnaistettu arviointitutkimus on viime vuosina ollut kasvussa. Hyvä esimerkki on Nesta-säätiön toteuttama *Creative Credits –ohjelman* arviointi satunnaistetulla koeasetelmalla Iso-Britanniassa. Ohjelman tarkoituksena oli auttaa pieniä ja keskisuuria (PK) yrityksiä verkostoitumaan luovien toimialojen yritysten kanssa hankkimalla niiltä asiantuntijapanoksia. Tulosten mukaan 4000 punnan tuki lisäsi yritysten innovaatioprojektien määrää ja liikevaihtoa kuuden kuukauden kuluttua toimenpiteestä, mutta tilastollisesti merkitseviä vaikutuksia ei enää havaittu vuoden jälkeen (Bakhshi ym., 2015).

Iso-Britannian Department for Business, Innovation, and Skills käynnisti tammikuussa 2014 satunnaistetun arviointikokeen kasvuseteli ohjelman (*Growth Voucher Programme*) vaikutusten arvioimiseksi (UK Department for Business, Innovation, and Skills, 2014). 2000 punnan arvoisen kasvusetelin tarkoituksena oli auttaa pieniä yrityksiä hankkimaan asiantuntijapanoksia ja arvioida näiden panosten vaikutuksia. Vaikutusten arviointi toteutetaan nelivuotisena satunnaistettuna arviointikokeena. Tutkimuksen loppuraporttia ja tilastollisia testejä vaikutuksista kiinnostuksen kohteena oleviin pääasiallisiin vastemuuttujiin ei vielä tätä kirjoitettaessa ole vielä saatavilla. Väliraportti kuukauden kohdalla havaittavista tuloksista viittaa siihen, että kolmasosa yrityksistä, joille oli osoitettu seteli, oli käyttänyt sitä (UK Department for Business, Innovation, and Skills, 2016). Tulokset viittasivat myös siihen, että seteli olisi lisännyt joitakin liiketoiminnan suunnittelun osa-alueita.

Kehittyvissä maissa yritysten toimintaa tehostavien neuvonta- ja konsultointi-interventioiden on havaittu satunnaistetuissa kenttäkokeissa lisäävän yritysten kannattavuutta ja työllisyyttä (Bloom ym. 2003 ja Bruhn ym. 2018). Yhdysvalloissa toteutetussa yli 4000 henkilöä kattaneessa GATE-koeasetelmassa arvioitiin yrittäjyyskoulutuksen (ts. (potentiaalisille) yrittäjille kohdistetun koulutuksen) vaikutuksia. Fairlien ym. (2015) mukaan yrittäjyyskoulutuksella oli vähäisiä vaikutuksia yrityksen kokoon ja yrittäjän kotitalouden tuloihin. Ainoa tutkijoiden löytämä vaikutus oli se, että kuusi kuukautta koulutuksen jälkeen koulutukseen osallistuneet omistivat yrityksen todennäköisemmin ja olivat todennäköisemmin töissä kuin verrokkiryhmään kuuluneet.

Kotimaassa yritystukia ei ole arvioitu satunnaisilla koeasetelmilla, mutta muilla politiikka-alueilla, kuten esimerkiksi työmarkkinoihin liittyen, niitä on tehty jonkin verran (ks. esimerkiksi Hämäläinen, Uusitalo ja Vuori 2008).¹ Yksi merkittävimmistä satunnaistetuista arviointikokeista on vuosina 2017-2018 toteutettu perustulokokeilu. Kokeilussa 2000:lle Kelan työttömyysetuutta saaneelle maksettiin 560 euron vastikkeetonta kuukausittaista perustuloa. Ensimmäisen vuoden tulosten mukaan perustulolla ei ollut vaikutusta työllisyyteen (Hämäläinen ym. 2019). Lisäksi verohallinto on tehnyt taloustieteilijöiden kanssa yhteistyötä ja hyödyntänyt satunnaistettuja koeasetelmia erilaisten toimien ja muutosten vaikuttavuutta selvittäessä (ks. mm. Kosonen ja Ropponen 2015 ja Hammar, Hyytinen ja Tuimala 2018).

¹ Yritystukia on kuitenkin jaettu joissakin tilanteissa satunnaisesti myös Suomessa.

Tässä raportissa esittelemme satunnaistettuja kontrolloituja kokeita yritystukien vaikutusarvioinnin näkökulmasta.² Kokeiden käytännön järjestelyihin liittyy lukuisia yksityiskohtia, jotka riippuvat ohjelmien tavoitteista, institutionaalisista erityispiirteistä ja tarvittavista aineistoista. Kokeiden suunnittelu edellyttää siksi aina tiivistä yhteistyötä tutkijoiden ja ohjelmantoteuttavien tahojen välillä. Eriteltyt koeasetelmat soveltuvat myös verotukien vaikutusten arvointiin. Verotukien osalta satunnaistaminen edellyttää kuitenkin lakimuutoksia, ja näin ollen lainsäädäntöprosessia ja yhteistyötä lakia valmistelevien tahojen kanssa. Sama koskee myös joitakin sellaisia suoria tuki, joista on säädetty laissa. Tuoreen esimerkin satunnaistetun koeasetelman lainsäädäntöprosessista antaa perustulokokeilu.

Kirjoituksen tarkoituksena on antaa lukijalle kattava mutta tiivis kokonaiskuva yritystukien satunnaistettuihin arviointitutkimuksiin liittyvistä keskeisistä käsitteistä ja kokeen onnistumisen kannalta oleellisista kysymyksistä. Yksi tämän raportin tavoitteista on, että se osaltaan lisää keskustelua, jota yritystukien vaikuttavuusarvioinneista päättävät, tukiaohjelmia toteuttavat tahot ja tutkijayhteisö aiheesta käyvät.

Raportti etenee seuraavasti: Seuraavassa luvussa (luku 2) esitellään lyhyesti satunnaiskokeisiin liittyvää käsitteistöä ja tilastollisia reunaehtoja. Luvussa 3 keskitymme eräisiin yritystukia koskeviin erityispiirteisiin ja -kysymyksiin. Näitä ovat mm. se, miten yritysten välinen mahdollinen vuorovaikutus tulisi koejärjestelyn suunnittelussa huomioida, miten yritysten omat (tuen tehokkuutteen vaikuttavat) päätökset tulee huomioida ja miten tukien vaikutusmekanismeja voitaisiin selvittää. Lisäksi tässä luvussa keskustellaan tuki-instrumenttien skaalautuvuudesta, vaikuttavuustulosten kontekstisidonnaisuudesta ja kokonaistaloudellisten vaikutusten arvioinnista. Luvussa 4 pohditaan lyhyesti satunnaistettua vaikuttavuusarviointia yritystukipolitiikan kehittämistyökaluna. Lopuksi luvussa 5 esitellään muutamia näkemyksiä siitä, miten satunnaistettuja kontrolloituja kokeita voitaisiin soveltaa Suomessa, kun halutaan ymmärtää yritystukien vaikuttavuutta. Lukija, joka haluaa vain tutustua johtopäätöksiimme, voi hypätä johdannon luettuaan suoraan tähän viimeiseen loppupohdintojamme ja johtopäätöksiämme esittelevään lukuun.

² Yksi yritystukien yhteydessä mainittu ongelma on, että osa yrityksistä ei välttämättä tiedä, millaisia tukia on saatavilla. Myös tämän ongelman laajuutta voitaisiin selvittää satunnaistettujen koejärjestelyiden avulla varsin helposti ja luotettavasti. Emme kuitenkaan käsittele tätä kysymystä tarkemmin tässä raportissa. Todettakoon kuitenkin se, että tukien saatavuuteen liittyvän tiedon satunnaistamista voidaan hyödyntää myös vaikuttavuusarvioinnissa.

2 Yleistä satunnaistetuista kontrolloiduista koejärjestelyistä

Satunnaistetun kontrolloidun koejärjestelyn tarkoituksena on selvittää, vaikuttaako toimenpide sen kohderyhmään. Vaikuttavuusarvioinnin keskeinen haaste on selvittää, mitä kohderyhmälle olisi käynyt ilman toimenpidettä. Tämä seuraa siitä, että toimenpiteelle altistettua ryhmäähän ei voida havaita samanaikaisesti tilanteessa, jossa heitä ei ole altistettu toimenpiteelle. Yritystukiin sovellettuna tämä tarkoittaa, että yritystukea saavaa yritysjoukkoa ei voida havaita samanaikaisesti tilanteessa, jossa nämä yritykset eivät olisi saaneet yritystukea.

Tässä osiossa annetaan kattava mutta tiivis kokonaiskuva yritystukiin satunnaistettuihin arviointitutkimuksiin liittyvistä keskeisistä käsitteistä ja kokeen onnistumisen kannalta keskeisistä kysymyksistä. Aihepiiriin yksityiskohtaista ja tilastoteoreettista esitystä etsivälle lukijalle on saatavilla lukuisia satunnaistettujen koejärjestelyiden tilastoteoriaa ja käytännön toteutusta käsitteleviä kattavia tekstejä. Tällaisia teoksia ja kirjoituksia ovat mm. Glennersterin ja Takavarashan (2013) kirja *Running randomized evaluations: A practical guide*, sekä Banerjeen ja Duflon (2017) toimittamat käsikirjat *Handbook of Field Experiments Volume 1 ja Volume 2*. Banerjeen ja Duflon käsikirjasta (Volume 1) löytyy mm. Atheyn ja Imbensin (2017) kirjoittama erinomainen katsausartikkeli *Econometrics of randomized experiments*.

2.1 Koe- ja verrokkiryhmä

Satunnaistamisen tärkein tehtävä on muodostaa verrokkiryhmä, johon toimenpidettä ei kohdisteta, mutta joka on keskimääräisiltä ominaisuuksiltaan samankaltainen kuin toimenpiteelle altistettava koeryhmä. Verrokkiryhmän tulema toimii toisin sanoen arviointiasetelmassa kontrafaktuaalina eli se kertoo, mitä niille yrityksille, jotka saivat yritystukea, olisi käynyt ilman toimenpidettä eli jos ne olisivat jääneet ilman tukea. Vertailukelpoisen verrokkiryhmän avulla voidaan huomioida se osa käyttäytymisen muutoksesta koeryhmässä, joka olisi toteutunut ilman tukeakin. Jos esimerkiksi liikevaihdon lisäystä tavoittelevaan tukiohjelmaan hakisi yrityksiä, jotka kasvattavat liikevaihtoaan myös ilman tukea, voidaan tämä tuesta riippumaton muutos ottaa huomioon, koska se havaitaan vertailukelpoisen ilman tukea jäävän verrokkiryhmän avulla.

Yksinkertaisessa, täydellisesti satunnaistetussa kontrolloidussa vertailukokeessa (*completely randomized experiment*) tarvittavat kaksi ryhmää saadaan muodostettua

satunnaistamalla kokeeseen valitut yritykset koe- ja verrokkiryhmään. Kokeeseen voidaan valita N :n yritystä, jotka yksinkertaisimmassa tapauksessa jaetaan kahteen yhtä suureen ($N/2$ suuruiseen) ryhmään. Toiseen ryhmään kohdistetaan toimenpide, kuten esimerkiksi yritystuki, joten se on koeryhmä. Toiseen ryhmään ei kohdisteta toimenpidettä: Jälkimmäinen ryhmä on siis verrokki- tai vertailuryhmä. Toinen esimerkki täydellisen satunnaistetusta koeasetelmasta on se, että kohderyhmään valitaan satunnaisesti N_t yritystä. Tällöin vertailuryhmään jää $N_c = N - N_t$ yritystä.

Koska ryhmät on satunnaisesti jaettu, voidaan niiden odottaa olevan keskimäärin samanlaisia. Tämä tarkoittaa, että vaikka yksittäiset yritykset ovat erilaisia, ovat satunnaistamalla muodostetut ryhmät havaittavilta ja ei-havaittavilta ominaisuuksiltaan keskimäärin samankaltaiset. Jos satunnaistaminen tehdään oikein, tämä ominaisuus (ts. ryhmien vertailukelpoisuus ja samankaltaisuus) pätee sitä suuremmalla todennäköisyydellä, mitä suurempi N on.

Tässä kohtaa on hyvä todeta se, että satunnaistaminen ei edellytä, että tukea saavat yritykset poimitaan satunnaisesti kaikkien yritysjoukosta. Koeasetelmassa on mahdollista ottaa huomioon esimerkiksi etukäteistieto siitä, miten hyvä yrityksen kehitys- tai investointisuunnitelma on (ks. myös alla keskustelu kohdepopulaation valinnasta).

Kuten alla kuvataan, satunnaistetun kokeen onnistumisen kannalta on oleellista, että sillä on riittävä tilastollinen voima. Tällöin vertailu- ja koeryhmiä ei välttämättä kannata muodostaa täydellisesti satunnaistamalla, vaan hyödyntämällä ns. jyvitettyä satunnaistamista (*stratified randomised experiment*; ks. Athey ja Imbens 2017, s. 99-109). Tätä menettelyä käsitellään alla lyhyesti lisää.

2.2 Kokeen koko

Kokeen koko eli kohde- ja verrokkiryhmiin satunnaistettavien yritysten lukumäärä (N) määrittää kokeen tarkkuuden ja luotettavuuden.

Mitä suurempi koe on, sitä tarkempia arvioinnin tulokset ovat tilastollisesti. Liian pieni kokeen koko johtaa siihen, että kohde- ja verrokkiryhmän välisiä eroja ei voida tilastollisesti testaamalla todentaa riittävän tarkasti. Tällä tarkoitetaan sitä, että vaikka koe- ja vertailuryhmien välillä havaittaisiin ero kiinnostuksen kohteena olevassa vastemuuttujassa, ei voida sulkea pois sitä, että kyse on sattumasta. Riskinä toisin sanoen on, että jos N on pieni, toimenpiteen vaikutus ei osoittaudu tilastollisessa testissä nollasta poikkeavaksi, vaikka sen vaikutus todellisuudessa olisi positiivinen tai negatiivinen. Kokeen koko täytyy suunnitella siksi etukäteen.

Kokeen tarvittava koko voidaan arvioida tilastoteorian avulla, kun tunnetaan mitattavien vastemuuttujien vaihtelu kohdepopulaatiossa. Mitä suurempi vastemuuttujien vaihtelu on (ts. mitä enemmän vastemuuttuja vaihtelee muista syistä), sitä suurempi koe tarvitaan, jotta toimenpiteen vaikutus saadaan mitattua riittävän tarkasti.

Jos keskeisistä vastemuuttujista ei ole valmista aineistoa, vaan ne muodostetaan kyselyllä kokeen aikana, voidaan kokeen kokoa arvioida selvittämällä vastemuuttujien vaihtelu etukäteiskyselyllä. Rekisteriaineistoista saatavien muuttujien käyttämisen yksi etu on se, että tällaisia etukäteiskyselyitä ei välttämättä tarvita.

Jos toimenpiteen vaikutuksen odotetaan olevan pieni, on kokeen koon on oltava suurempi kuin tilanteessa, jossa toimenpiteen vaikutuksen odotetaan olevan suuri. Tämä seuraa siitä, että pienen vaikutuksen tilastollinen todentaminen edellyttää suurempaa tilastollista tarkkuutta. On myös hyvä todeta tässä se, että tuen myöntäjä ja satunnaisesti kokeen järjestäjä voivat vaikuttaa vaikutuksen odotettuun suuruuteen. Esimerkiksi yritystuen kokoa voidaan kasvattaa.

Kokeen onnistumisen kannalta on tärkeää, että kokeessa on mukana riittävä määrä yrityksiä ja toimenpide on niin suuri, että sillä voidaan odottaa olevan tilastollisesti tunnistettavia vaikutuksia. Tällöin sanotaan, että satunnaisella arviointikokeella on riittävä tilastollinen voima. Tilastoteoreettisesti riittävällä voimalla tarkoitetaan hieman yksinkertaisten sitä, että vaikutuksiltaan nollassa poikkeava toimenpide todetaan vähintään 95% toistettavista kokeista vaikuttavaksi.

Kun tilastollisen testin voima kasvaa, todennäköisyys tehdä ns. tyyppin II päättelyvirhe pienenee. Tyyppin II päättelyvirhe tapahtuu, jos toimenpiteen (yritystuen) vaikutus ei osoittaudu tilastollisessa testissä nollassa poikkeavaksi, vaikka sen vaikutus todellisuudessa olisi positiivinen (tai negatiivinen).

Kuten yllä sanotusta voi päätellä, riittävän suuri koe- ja vertailuaineisto on konteksti- ja tukisidonnainen: Tarkoituksenmukainen N riippuu sekä siitä, minkälainen arviointitilanne kohdataan että siitä, minkä verran valituissa tulemamuuuttujissa on vaihtelua. Tämä tarkoittaa, että yksiselitteistä vastausta siihen, mikä on riittävän iso N, jotta satunnaistamista voidaan tehdä, ei ole.

2.3 Tutkimuskysymyksen ja -hypoteesien asettaminen

Kokeensuunnittelun tärkeä osa on tutkimuskysymysten muodostaminen ja tutkimushypoteesien asettaminen. Tutkimuskysymyksiä ohjaa kokeen kohteena olevan toimenpiteen tavoitteet ja perustelut.

- Jos toimenpiteen tavoitteena on esimerkiksi tuetun yrityksen liikevaihdon lisääminen, on tarkoituksenmukainen tutkimuskysymys: ”Lisäkö toimenpide tuettujen yritysten liikevaihtoa?” Tällöin tilastollisessa analyysissä testataan nollahypoteesia, että ohjelmalla ei ole vaikutusta liikevaihtoon. Jos nollahypoteesi hylätään, johtopäätös on, että toimenpide vaikutti liikevaihtoon. Tarkempi tarkastelu sitten paljastaa, kasvoiko vai supistuiko liikevaihto.
- Jos toimenpiteen tavoitteena on lisätä liikevaihtoa koko talouden tasolla (kaikkien yritysten yhteenlaskettu liikevaihto), on lisäksi tarkoituksenmukaista tarkastella mahdollisia negatiivisia (tai positiivisia) markkinavaikutuksia tukea ilman jääviin yrityksiin. Käsittelemme tällaisten välillisten tai epäsuorien vaikutusten arviointia yksityiskohtaisemmin jäljempänä.

Tilastollisessa testaamisessa löydetään usein tilastollisesti merkitsevä tulos, kun tarkastellaan riittävän useata vaihtoehtoista vastetta. Tämä johtuu siitä, että satunnaisesta vaihtelusta johtuen jokin lukuisasta vasteesta antaa todennäköisesti merkitsevän tuloksen. Tutkimuksen uskottavuuden lisäämiseksi on usein hyödyllistä määritellä pääasialliset tutkimushypoteesit, joita testataan yhdellä tai muutamalla pääasiallisilla vastemuuttujilla (Olken, 2015).

Pääasialliseen hypoteesiin on usein tarkoituksenmukaisinta sisällyttää vain muutaman vastemuuttujan testaaminen. Syy on se, että testattavien vastemuuttujien lisääminen pienentää tyypillisesti kokeen tilastollista voimaa jyrkästi, jos tilastollisissa testeissä otetaan huomioon usean hypoteesin samanaikainen testaaminen. Pääasiallisena muuttujana voidaan myös käyttää useammasta ohjelman tavoitetta mittaavasta vasteesta muodostettua indeksiä.

Jos ensisijaista nollahypoteesia, että kokeella ei ole vaikutusta ensisijaisiin vasteisiin, ei hylätä, on kokeesta vedettävä johtopäätös se, että toimenpiteelle ei löydetty vaikutusta.

Ensisijaisen hypoteesin lisäksi voidaan määritellä toissijaisia hypoteeseja. Nämä voivat olla hypoteeseja suhteessa muissa tutkimuksissa käytettyihin vasteisiin, joka parantaa tulosten vertailtavuutta. Voi olla myös hyödyllistä tarkastella toissijaisia muuttujia, jotka

voivat valaista mekanismeja toimenpiteen vaikutusten taustalla. Vaikka toissijaisten hypoteesien testaaminen antaisi tilastollisesti merkitseviä tuloksia, ajatus on, että ne eivät voi kumota ensisijaisten hypoteesin tuottamaa johtopäätöstä.³ Tätä valaisee seuraava esimerkki:

- Ensisijaisena hypoteesina on ennen koetta määritelty testattavaksi tuen vaikutus liikevaihtoon. Tilastollisessa testissä ei hylätä ensisijaista hypoteesiä, että tuella ei ole vaikutusta liikevaihtoon.
- Toissijaisena hypoteesina on ennen koetta määritelty testattavaksi tuen vaikutus työllisyyteen. Tilastollisessa testissä hylätään nollahypoteesi, että tuella ei ole vaikutusta työllisyyteen.
- Vaikka toissijainen hypoteesi viittaa työllisyysvaikutuksiin, on kokeen johtopäätöksenä se, että tuella ei ollut tilastollisesti merkitsevää vaikutusta.

Edellisessä esimerkissä on tärkeää muistaa se, että ensisijaiseksi hypoteesiksi olisi voitu myös määritellä ennen koetta työllisyysvaikutuksen testaaminen, tai työllisyys- ja liikevaihtovaikutusten testaaminen. Jälkimmäisessä testissä tulee huomioida usean samanaikaisen hypoteesin testaaminen, joka voi johtaa siihen, että vaikka yksittäinen hypoteesi hylätään (tässä työllisyyden kohdalla), ei yhteistestaaminen välttämättä tuota merkitsevää tulosta.

Seuraava esimerkki pyrkii valaisemaan, mitä ensisijaisilla ja toissijaisilla hypoteeseilla tarkoitetaan, kun tutkitaan esimerkiksi innovaatiotukien vaikuttavuutta:

- *Ensisijainen hypoteesi:* Jos tuen ensisijaisena tavoitteena on lisätä innovaatio-toimintaa, tulee ensisijaisena vastemuuttujana käyttää sitä mittaavaa muuttujaa. Sellaisia ovat mm. yritysten T&K-investoinnit, yrityksissä tehtyjen T&K-työtuntien lukumäärä, yrityksissä tehtyjen tuotantoprosessi- ja tuoteinnovaatioiden lukumäärä, yritysten patenttien lukumäärä jne. Sopivin ensisijainen vastemuuttuja voi riippua lukuisista tekijöistä, kuten ohjelman erityisluonteesta (esim. erityisteemasta, kuten yliopistoyhteistyö tai energiatehokkuus), tuen kohdentamisesta ja muuttujien vaihtelun suuruudesta kohdepopulaatiossa. Myös arvoinnin aikajänne on ensisijaisen vastemuuttujan valinnassa oleellinen. Mikäli aikajänne on lyhyt, tulee valita sellainen muuttuja, jossa voidaan odottaa nähtävän nopeita vaikutuksia.

³ Kuten tässä raportissa kuvataan myöhemmin, yritystukien vaikutusmekanismeja voidaan pyrkiä tutkimaan tarkemmin myös suunnittelemalla satunnaistettu koeasetelma erityisellä tavalla.

- T&K-investoinnit ovat saatavilla laajalle yrityspopulaatiolle Tilastokeskuksen T&K-kyselystä ja niiden voidaan useissa tapauksissa odottaa muuttuvan melko nopeasti tuen myötä. Nämä ovat sen käyttöä puoltavat tekijät. T&K-kyselystä on saatavilla myös yritysten innovaatioaktiivisuutta kuvaavia tekijöitä kuten tieto siitä, tekikö yritys prosessi- tai tuoteinnovaatioita.
 - Koko yrityspopulaation kattava innovaatiotoimintaa kuvaava tieto on saatavilla julkisten patenttihakemusten ja -myöntöjen tietokannoista. Patenttihakemukset tehdään keskimäärin melko nopeasti T&K-projektia toteutettaessa (kirjallisuus alkaen Hall, Griliches ja Hausmanista (1986)), joten niillä voidaan päästä kiinni melko nopeisiin vaikutuksiin. Toisaalta ne mittaavat vain osaa yritysten innovaatiotoiminnasta, eivätkä ne kata esimerkiksi kaikkia sellaisia tuotantoprosesseihin liittyviä innovaatioita, joita yritykset eivät halua paljastaa kilpailijoille.
 - Mikäli ohjelman tavoitteena on lisätä tuetun yrityksen liikevaihtoa, on se ilmeinen ensisijainen vastemuuttuja. Sen kohdalla on tärkeää huomoida mm. se, että tukien vaikutus liikevaihtoon voi syntyä viiveellä (ks. mm. Einiö 2014)
 - Mikäli ohjelman kannalta sopivaa ensisijaista vastetta ei ole saatavilla olemassa olevista aineistoista, voidaan se muodostaa kyselyllä.
- *Toissijainen hypoteesi:* Toissijaiseksi hypoteesiksi voidaan valita vaihtoehtoisten vastemuuttujien testaamista, mikä auttaa ymmärtämään pääasiallisessa vastemuuttujassa tapahtuvien muutosten luonnetta ja taustalla vaikuttavia mekanismeja, sekä voi lisätä tulosten vertailtavuutta aiempiin tutkimuksiin. Innovaatiotukiesimerkissä toissijaisiin vaisteisiin voidaan sisällyttää mm. muita innovaatiotoimintaa mittaavia vasteita tai vaikkapa kansainvälisiltä markkinoilla syntyvän liikevaihdon määrää. Lisäksi voidaan tarkastella mm. yrityksen työvoimassa ja muissa panoksissa tapahtuvia muutoksia tai sitä, avaako yritys uusia toimipisteitä ulkomailla.

Tutkimushypoteesit voivat sisältää myös osapopulaatioiden tarkasteluita. Osapopulaatioita tarkasteltaessa testaaminen on myös sopeutettava testattavien hypoteesien lukumäärään. Lisäksi osapopulaatiotarkasteluissa on käytettävissä vain osa kokeeseen osallistujista. Näistä seuraa se, että osapopulaatiotarkasteluiden lisääminen pääasiallisiin tutkimushypoteeseihin edellyttää suurempia otoskokoja. Mikäli vaikutukset osapopulaatioissa eivät ole tutkimuskysymyksen kannalta ensisijaisia, voidaan niitä koskeva testaaminen sisällyttää toissijaisiin hypoteeseihin.

Seuraava esimerkki valaisee, mitä osapopulaatioiden tarkastelu tarkoittaa, kun arvioidaan innovaatiotukien vaikuttavuutta:

- *Ensisijainen hypoteesi:* Innovaatiotukien vaikutuksista voidaan olla ensisijaisen kiinnostuneita innovaatiointensiivisillä toimialoilla, mikäli tällaisilla toimialoilla tuen odotetaan vaikuttavan eniten. Ensisijainen hypoteesi voidaan tällöin määritellä testattavaksi esimerkiksi toimialoilla, joilla ennen koetta havaitaan toteutettavan eniten innovaatioita suhteessa yritysten lukumäärään, työllisyyteen tai liikevaihtoon.
- *Toissijainen hypoteesi:* Eräs erityisryhmä, jossa innovaatiotukien voidaan odottaa vaikuttavan voimakkaimmin, ovat rahoitusrajoitteista kärsivät yritykset (ks myös Hyytinen ja Toivanen 2005). Ulkoisen rahoituksen heikko saatavuus voi olla erityinen este korkean riskin T&K-investoinnille. Tuen vaikutusta voi tällöin olla kiinnostavaa ja tarkoituksenmukaista testata esimerkiksi sellaisessa yritysryhmässä, jonka kassatilanne (joustovara omassa rahoituksessa) on suhteellisen heikko.

Koska taloustieteellisessä tutkimuksessa mekanismien tutkimisella on keskeisempi rooli kuin monella muulla tieteenalalla, ja ne muotoutuvat usein osittain myös ensimmäisten tutkimustulosten perusteella, voi joskus olla perusteltua jättää toissijaisten hypoteesien täsmällinen määrittely etukäteisuunnitelmassa melko yleiselle tasolle, ja määrittää toissijaisia vasteita ja otostarkasteluita tutkimuksen edetessä (Olken, 2015).

2.4 Kohdepopulaation valinta

Laajin mahdollinen kohdepopulaatio satunnaistetulle vaikuttavuusarvioinnille on koko yrityssektori. Tällöin koeryhmään eli tuen saajiksi valitaan satunnaisesti yrityksiä koko yritysryhmästä ja verrokkiryhmänä toimii kokeen ulkopuolelle jäävät yritykset. Kuten alla toteamme, tämä ei luonnollisesti ole useinkaan tarkoituksenmukaista, koska esimerkiksi innovaatiotukia ei ole perusteltua kohdistaa yrityksille, joilla ei ole innovointimahdollisuuksia tai -tavoitteita.

Kokeen tilastollisen voiman lisäämiseksi käytetään usein jyvitettyä satunnaistamista (*stratified randomized experiment*), kun käytössä on laaja kohdepopulaatio. Tällöin satunnaistaminen tehdään osajoukkojen sisällä, kuten esimerkiksi toimialoittain tai sen jälkeen, kun yritykset on ensin jyvitetty ryhmiin niiden koon mukaan. Jyvittämistä käytetään myös silloin, kun halutaan selvittää vaikutuksia eri osapopulaatioissa.

Kohdepopulaation rajaaminen voi olla tarpeellista useista syistä. Toimenpiteitä, joiden tarkoitus on vaikuttaa tiettyihin yritysryhmiin, esimerkiksi toimialoittain tai koon mukaan, voi olla perusteltua arvioida rajatuissa ryhmissä. Kohdepopulaatiota voidaan rajata

myös siksi, että toimenpiteen vaikutuksia halutaan tutkia vain yritysjoukossa, joka haakeutuu ohjelman piiriin. Tämä rajausta jättää ulkopuolelle yritykset, jotka eivät hae ohjelmaan. Nämä yritykset voivat jättäytyä ohjelman ulkopuolelle esimerkiksi siksi, että ne huomioivat myöntäjän valitsemiskriteerit ja pitävät myönteistä päätöstä epätodennäköisenä. Tämä ei kuitenkaan suoraan tarkoita sitä, että vaikutukset tässä ryhmässä ovat pieniä, ja toimenpiteen oikean kohdentamisen kannalta voisi olla hyödyllistä arvioida tuen vaikutusta myös niiden joukossa.

Kohderyhmän rajaaminen rajoittaa tutkimustulosten yleistettävyyttä. Käsittelemme yleistettävyyttä tarkemmin alla.

2.5 Satunnaistaminen käytännössä

Satunnaistaminen voidaan toteuttaa monella eri tavalla. Sopiva satunnaistamismenetely riippuu arvioitavasta toimenpiteestä.

- Kun toimenpide kohdistetaan yrityspopulaation, jota ei ehdollisteta tuen (ohjelmaan) hakemiselle, voidaan satunnaistaminen toteuttaa koepopulaatiolle yhdellä kerralla käyttämällä esimerkiksi kohdepopulaation kattavaa rekisteriaineistoa.
- Satunnaistaminen voidaan toteuttaa yhdellä kerralla myös silloin kun kohdepopulaatio rajataan ohjelmaan hakeneisiin, jos ohjelmaan haulla on yhteinen takaraja. Satunnaistaminen yhdellä kerralla ei ole mahdollista, jos kohdepopulaatio rajataan ohjelmaan hakeviin ja yritykset osallistuvat toimenpiteeseen eriaikaisesti. Tällöin satunnaistaminen voidaan toteuttaa esimerkiksi hakuajan suhteen. Mikäli ohjelma on sen luonteinen, että yritykset hakevat siihen jatkuvasti, ei kertaluonteinen koe- ja verrokkiryhmän muodostaminen ole välttämättä mahdollista. Tällöin yrityksiä voidaan esimerkiksi kerätä hakemisjärjestyksessä ryppäiksi ja satunnaistaminen voidaan tehdä näin muodostetussa pienemmissä ryhmissä. Tilastollisessa analyysissä lopulliset koe- ja verrokkiryhmät muodostetaan tällöin osaryhmät yhdistämällä. Toisena vaihtoehtona tämän tyyppisessä tilanteessa voi olla yritysten satunnaistaminen hakujärjestyksen mukaan, jolloin etukäteen määritetään satunnaisesti niiden hakijoiden hakujärjestysnumero, jotka valitaan koeryhmään.
- Mikäli olemassa olevaan toimenpiteeseen eivät osallistu kaikki yritykset, jotka siihen voisivat osallistua, voidaan satunnaistaa kannustimia osallistua ohjelmaan. Satunnaisesti valitulle ryhmälle voidaan esimerkiksi lähettää tietoa ohjelman hyödyistä ja tarjota apua ohjelmaan osallistumiseen vaadittavan hakemuksen tekemisessä. Tämän menettelyn heikkous on se, että se ei anna täyttä kontrollia koeryhmän suuruudesta, jolloin kokeen tilastollinen voima voi muodostua liian pieneksi.

- Kokeen kohderyhmä voi joskus olla tarkoituksenmukaista määritellä myös jossakin ohjelmaan hakeneiden yritysten osapopulaatiossa. Jos ohjelmaan hakee enemmän yrityksiä, kuin siihen voidaan valita, voidaan satunnaisesti valita hakeneiden joukosta ne yritykset, jotka eivät pääse mukaan. Tällöin saadaan selvitettyä toimenpiteen vaikutus ohjelmaan hakevassa yrityspopulaatiossa. Tässä menettelyssä selvityksen ulkopuolelle jäävät yritykset, jotka eivät hae ohjelmaan, ja tämä rajausta heikentää tulosten yleistettävyyttä.
- Joissakin tapauksissa satunnaistaminen voi olla tarkoituksenmukaista toteuttaa ohjelman hyväksymisrajalla.⁴ Hyväksymisraja voi muodostua esimerkiksi siten, että hakemukset pisteytetään, ja ohjelmaan otetaan mukaan vain tietyn pisterajan ylittävät hakijat. Satunnaistettu arviointikoe voidaan toteuttaa yritysjoukossa, joka jää rajan alapuolelle. Rajalla satunnaistamisen etuna on se, että se mahdollistaa tuen kohdentamisen suurelta osin tuen jakajan linjausten mukaisesti. Sen heikkous on kohdepopulaation rajaaminen osaan yrityksiä (rajan alapuolelle jäävät). Tämä heikentää tulosten yleistettävyyttä. Tästä syystä julkisten toimijoiden toteuttamien ohjelmien kohdalla on suositeltavampaa käyttää satunnaistamismenettelyä, joka sisällyttää kokeeseen laajemman kohdejoukon. Yleistettävyyttä voidaan parantaa rajaa nostamalla koe- ja verrokkiryhmän muodostamisen ajaksi, jolloin mukaan sadaan laajempi yritysjoukko. Yritystukiarviointia laajemmassa kontekstissa rajalla satunnaistaminen voi olla hyödyllinen menettely tietyissä erityisemmissä koejärjestelyissä, joissa esimerkiksi yritysrahoitusta ei jaa julkinen taho, vaan yksityisen sektorin toimijoita kuten rahoituslaitokset, joiden kannustimet satunnaistamiseen laajemmassa yritysjoukossa voivat olla rajalliset.

Yritystukiohjelmien vaikutukset ovat usein pitkäaikaisia ja saattavat ilmetä viiveellä. Tästä syystä jotkin satunnaistamismenettelyt, joita käytetään muissa konteksteissa, eivät usein ole yritystukiohjelmien arviointiin sopivia. Mm. yritysten nopea kiertäminen (*rotation*) ohjelmassa, jolloin yritykset kiertävät koe- ja verrokkiryhmän välillä suhteellisen tiheään tahtiin, tai osallistumisen tiivis jaksottaminen (*phasing-in*), jossa verrokkiryhmänä toimii hiukan myöhemmin toimenpiteen piiriin tulevat yritykset, ovat harvoin sopivia menetelmiä, koska ne edellyttävät toimenpiteen suhteellisen nopeata vaikutusta.

Satunnaistamisen taso on yritystukiohjelmissa tyypillisesti yritys, mutta myös muut tasot voivat olla tarkoituksenmukaisia. Esimerkiksi yritysten neuvontapalveluja arvioitaessa voidaan tietyissä tapauksissa satunnaistaminen toteuttaa neuvojatasolla, jos neuvojien lukumäärä on riittävän suuri tavoitellun tilastollisen voiman tuottamiseksi. Kon-

⁴ Rajalla satunnaistaminen on satunnaistetun koejärjestelyn muoto, joka ei ole sama kuin jäljempänä käsiteltävä regressioepäjatkuvuus-asetelma, jossa satunnaistamista ei aktiivisesti tuoteta osana koejärjestelyä.

sernien tai monitoimipaikkaisten yritysten sisäisten epäsuorien vaikutusten aiheuttaman estimointiharhan kontrolloimiseksi satunnaistaminen voidaan tehdä konserni- tai yritystasolla (voi esimerkiksi olla niin, että tuki myönnetään verovelvollisen yritysyksikön tasolla, mutta käytettävissä on toimipaikkatason aineisto, kuten Laukkasen ym (2019) tutkimuksessa energiaveropalautuksen vaikutuksista).⁵ Epäsuoria vaikutuksia käsitellään tarkemmin alla.

2.6 Satunnaistettava toimenpide

Koetta suunniteltaessa on valittava satunnaistettava toimenpide. Valintaan vaikuttaa tutkimuskysymys ja tarvittava tilastollinen voima.

Satunnaistettava toimenpide voi olla ohjelmaan osallistumisen tai tuen intensiteetin (esim. tukitason) satunnaistaminen. Näillä menettelyillä voidaan vastata eri kysymyksiin:

- Tukiohjelmaan osallistumisen satunnaistaminen avulla voidaan selvittää ohjelmaan osallistumisen keskimääräinen vaikutus valitulla tukitasolla. Satunnaistaminen voidaan tällöin toteuttaa sopivimmaksi katsotussa kohdepopulaatiossa. Jos kohdepopulaatioksi valitaan koko yrityssektori, saadaan toimenpiteen keskimääräinen vaikutus (*average treatment effect*). Keskimääräistä vaikutusta voidaan myös arvioida osapopulaatioissa, mutta tämä on syytä kirjata etukäteen koesuunnitelmaan ja laskea kokeen tilastollinen voima osatarkastelut huomioiden.
- Toinen keskeinen vaihtelu toimenpiteessä on tuen intensiteetin vaihtelu. Esimerkiksi suorien tukien kohdalla asetelma vastaa kysymykseen siitä, kuinka paljon tukieuron lisääminen vaikuttaa tuettujen yritysten joukossa. Myös esimerkiksi neuvontapalveluiden ajallista kestoja tai sisältöä voidaan vaihdella. Koska tässä asetelmassa satunnaistetaan tuen määrää (intensiteettiä), tarkastelu edellyttää, että kaikki kokeeseen osallistuvat yritykset ovat mukana tukiohjelmassa. Julkisten toimijoiden ohjelmien yhteydessä on kuitenkin perusteltua tarkastella yleisempiä vaikutuksia laajemmassa kohdejoukossa kuin tuetut yritykset: Jos vain intensiteettiä vaihdellaan ja tarkastellaan vain tukea saavia yrityksiä, ei tukien epäsuoria tai välillisiä vaikutuksia useinkaan voida selvittää.

⁵ Tällaisissa tilanteissa voi olla useita tapoja selvittää, onko tuloksissa harhaa epäsuorien vaikutusten vuoksi. Esimerkiksi näennäiskoeasetelmissa voidaan tarkastella yksitoimipaikkaisten yritysten osaotosta, jossa saman yrityksen (tai verovelvollisen yritysyksikön) toimipaikkojen välisiä epäsuoria vaikutuksia ei luonnollisestikaan ole, koska otoksen yrityksillä on vain yksi toimipaikka.

2.7 Vastemuuttajat: Rekisteriaineistot vs. tietojen kerääminen kyselyillä

Suomessa on kansainvälisessä mittakaavassa poikkeuksellisen hyvät rekisteriaineistot. Ne helpottavat satunnaistettujen arviointikokeiden toteuttamista tarjoamalla koko yrityspopulaatiolle taustatietoja ja monessa tapauksessa myös arvioinnin kannalta hyödyllisiä vasteita. Rekisteriperäisten aineistojen etuna on myös se, että ne tuottavat pitkäikäisiä aineistoja, joissa olemassa olevat yritykset havaitaan jo ennen koetta. Tämä auttaa arvioimaan tarvittavaa otoskokoa riittävän tilastollisen voiman tuottamiseksi. Lisäksi muuttujia koetta edeltävältä jaksolta voidaan käyttää koe- ja verrokkiryhmien samankaltaisuuden testaamiseksi (*balance test*).

Taulukko 1. Satunnaistetun arviointikokeen etukäteissuunnitelman muistilista (Olken, 2015).

Kohta	Lyhyt kuvaus
Pääasiallinen vastemuuttaja(t)	Tutkimuksen ensisijainen kiinnostuksen kohteena oleva vastemuuttaja. Jos tutkitaan useita vastemuuttujia, tulee hypoteesien testaamisessa huomioida usean samanaikaisen hypoteesin testaaminen (multiple hypothesis testing).
Toissijaiset vastemuuttajat	Muut kiinnostuksen kohteena olevat vastemuuttajat.
Muuttujien määrittelyt	Täsmällinen muuttujien määrittely, jossa esitetään miten alkuperäinen aineisto muunnetaan varsinaiseksi tutkimuksessa käytettäväksi muuttujiksi.
Havaintojen sisällyttämis-/poistamisäännöt	Havaintojen sisällyttämisessä ja poistamisessa käytettävät säännöt sekä säännöt puuttuvien havaintojen käsittelylle.
Tilastollisten mallien määrittely	Käytettävän tilastollisen mallin täsmällinen määrittely ja testattavat hypoteesit.
Kontrollimuuttajat	Käytettävien kontrollimuuttujien listaaminen ja niiden käytön määrittely
Tarkasteltavat osaotokset	Osaotoksilla tehtävän heterogeenisuus analyysin määrittely
Muuta	Muita suunnitelman osia, kuten aineiston monitorointisuunnitelma, kokeen pysäyttämissäännöt ja aineoston välivaiheen analysoinnin suunnitelma.

Rekisteriaineistot tarjoavat laajan aineistopohjan, mutta eivät aina sisällä ohjelman tavoitteiden toteutumista mittaavia tai tutkimuskysymykseen sopivia muuttujia. Tällöin muuttujat on tuotettava kyselytutkimuksilla, ja tilastollisen voiman arvioimiseksi empiiristen hajontalukujen avulla on tehtävä etukäteiskysely. Kyselytutkimusten heikkoutena on vastaajakato, joka voi aiheuttaa harhaa estimoinneissa ja tulee huomioida. Toinen

kyselytutkimusten heikkous on vastaamisharha, jolla tarkoitetaan vastaajien virheellisesti antamaa tietoa joko tarkoituksellisesti tai tarkoituksettomasti. Satunnaisille arviointikokeille erityinen ongelma voi olla tarkoituksellinen virheellisesti vastaaminen, joka on systemaattisesti erilaista koe- ja verrokkiryhmissä. Toimenpiteen saavat yritykset voivat esimerkiksi liioitella toimenpiteen hyötyjä, koska ajattelevat sen lisäävän todennäköisyyttä hyötyä toimenpiteestä myös tulevaisuudessa.

Eriyinen yritysryhmä, jota ei havaita ennen koetta rekisteriaineistoissa, ovat aloittavat yritykset. Näiden yritysten kohdalla on käytettävä epäsuoraa menettelyä arvioitaessa tilastollista voimaa empiiristen hajontalukujen avulla. Tällöin voidaan käyttää mahdollisimman hyvin kohderyhmää vastaavien juuri aloittaneiden yritysten muuttujia. Kokeeseen osallistuvilta aloittavilta yrityksiltä tulisi myös kartoittaa etukäteiskyselyllä taustatietoja, joiden avulla voidaan testata koe- ja verrokkiryhmien samankaltaisuutta.

2.8 Heterogeeniset vaikutukset ja tulosten yleistettävyys

Toimenpiteen vaikutukset voivat olla erilaisia eri kohdepopulaatioissa. Jossakin yritysjoukossa tuella voi olla suuria vaikutuksia kun taas toisessa pieniä. Tällöin vaikutuksia kutsutaan heterogeenisiksi (*heterogeneous effects*).

Tieto vaikutusten heterogeenisuudesta auttaa kohdentamaan tukia tehokkaasti niihin yrityksiin, joita tukemalla saadaan aikaiseksi suurimmat vaikutukset. Toisaalta ne rajoittavat tulosten yleistettävyttä, kun satunnaistettu arviointitutkimus yhdessä osapopulaatioissa ei kerro vaikutuksia muissa osapopulaatioissa. Näistä syistä tukien vaikutuksia on hyödyllistä tutkia mahdollisimman monissa osapopulaatioissa.

2.9 Arvioinnin avoimuus ja toistettavuus

Hyvän arviointitutkimuksen tulee olla avoin ja toistettava. Avoimuudella tarkoitetaan tutkimuksessa tuotettujen tulosten ja aineistodokumentaatioiden julkisuutta. Tämä sisältää myös sen, että tutkimusraporteissa selostetaan täsmällisesti ja tarkasti tutkimuksessa käytetyt tilastolliset menetelmät ja mallit ja aineiston käsittely. Hyvän käytännön mukaista on myös julkaista aineiston käsittelyssä ja tutkimustulosten tuottamisessa käytetyt ohjelmasyntaksit, mukaan lukien satunnaistamisessa käytetty ohjelmallinen protokolla.

Tutkimuksen täsmällinen toistettavuus edellyttää tutkimusaineiston julkaisemista tai sitä, että se on muiden tutkijoiden saatavilla jonkinlaisen erityisjärjestelyn avulla. Tällöin muut tutkijat voivat toistaa tutkimuksen tulokset.

On mahdollista, että tutkimusaineisto on luonteeltaan sellainen, että sitä ei voida luovuttaa muiden ulkopuolisten tutkijoiden käyttöön edes erityisjärjestelyin. Jos kuitenkin tutkimuksen raportoinnissa on noudatettu täyttä avoimuutta, voidaan tutkimuksen tulokset verifioida uudella vastaavan kaltaisella kokeella.

2.10 Arviointisuunnitelman etukäteisrekisteröinti

Kuluttajamarkkinoille hakevien lääkkeiden vaikutusten arviointeja koskevat tiukat etukäteisrekisteröintivelvoitteet, joissa kokeet kuvataan seikkaperäisesti. Tällä pyritään estämään tulosten manipulointi, mikäli ne eivät muuten tue lääkkeen markkinoille tuomista. Yhteiskuntatieteellisessä tutkimuksessa kannustimet manipulointiin ovat pienemmät, mutta toivottujen tulosten tuottaminen esimerkiksi lukuisia vaihtoehtoisia vartemuuttujia läpikäymällä (*data mining*) on silti varteenotettava riski. Tähän ongelmaan yksinkertainen ratkaisu on etukäteissuunnitelman tallentaminen arkistoon. Muun muassa *American Economic Association* ylläpitää arkistoa, johon tutkijat voivat tallettaa etukäteissuunnitelman ennen koetta ja määrittää sen julkaistavaksi tutkimusraportin yhteydessä.

3 Erityiskysymyksiä

Tässä luvussa käsitellään muutamia erityiskysymyksiä, joihin on tarpeen kiinnittää huomiota, kun yritystukien vaikuttavuutta arvioidaan satunnaistettujen vertailukokeiden avulla.

3.1 Erityiskysymys 1: Ulkoisvaikutukset

3.1.1 Miksi ulkoisvaikutuksiin pitää kiinnittää huomiota?

Tyypillisesti yritykset eivät toimi tyhjiössä, vaan ne ovat vuorovaikutuksessa muiden yritysten kanssa. Tämä vuorovaikutus voi liittyä suoraan yhteydenpitoon, kuten mm. siihen että yrityksellä on erilaisia toimittaja- ja alihankkijaverkostoja ja muita sopimussuhteita. Toisinaan vuorovaikutus on epäsuoraa, kuten silloin kun yritykset ovat vuorovaikutuksessa markkinoiden välityksellä: Keskenään kilpailevat yritykset reagoivat toistensa valintoihin strategisen vuorovaikutuksen vuoksi. Lisäksi vuorovaikutus voi olla tahatonta. Näin tapahtuu esimerkiksi jos yritys oppii uusia toimintamalleja tai teknologisia ratkaisuja kilpailijoiltaan tai toimialan muilta yrityksiltä.

Tämän vuoksi on syytä pohtia mahdollisuutta, että tietyn yrityksen saama yritystuki voi vaikuttaa myös muihin yrityksiin, vaikka nämä muut yritykset eivät itse saisikaan yritystukea. Taloustieteellinen tutkimus korostaa sitä, että erityisesti innovaatiotukia osin myönnetään juuri siitä syystä, että ne vaikuttavat myös muiden kuin tukea saavien yritysten innovaatio toimintaan tai tuottavuuteen. Yhden yrityksen tutkimus- ja kehitystoiminnan tuloksista saattavat nimittäin hyötyä muutkin yritykset. Jos näin käy, yritysten tutkimus- ja kehitystoiminnan voidaan sanoa aiheuttavan ulkoisvaikutuksia. Tämä tarkoittaa, että on mahdollista, että esimerkkiyrityksen A saama innovaatiotuki voi vaikuttaa myös yritysten B ja C innovaatio toimintaan tai tuottavuuteen, vaikka ne eivät itse saisi ko. tukea.

Toisaalta yritystuki voi vaikuttaa epäsuorasti markkinamekanismien välityksellä samoilla tuote- tai tuotannontekijämarkkinoilla kilpailevien yritysten keskuudessa. Esimerkiksi kone- tai T&K-investointeihin tukea saava yritys saa kilpailuetua ilman tukea jääviin mutta samalla markkinalla toimiviin yrityksiin nähden. Yksi tuen saamiseen liittyvä ja kilpailuasetelmaa (mahdollisesti) muuttava seuraus on se, että tukea saavan yrityksen kulut pienenevät. Tämän turvin tukea saava yritys voi esimerkiksi kasvattaa mark-

kinaosuuttaan tuotteiden hintoja laskemalla. Tästä seuraa epäsuora vaikutus tukea ilman jääviin yrityksiin, joka voi ilmetä esimerkiksi niiden markkinaosuuden pienentymisenä. Myös erilaisilla työllistämistuilla voi olla vastaavankaltaisia heijastusvaikutuksia: Kun joukko yrityksiä lisää rekrytointejaan jonkin työllistämistä lisäävän tuen vuoksi, työmarkkinoille jää vähemmän vapaata työvoimaa. Työllistämistuet voivat vaikuttaa työvoiman kohdentumiseen eri yritysten välillä erityisesti, jos työmarkkinat ovat paikalliset tai segmentoituneet ja jos työn tarjonta on melko joustamatonta. Taloustieteessä yritystukien markkinatason vaikutuksia kutsutaan usein yleisen tasapainon vaikutuksiksi (*general equilibrium effects*).

Edellä kuvatusta yritysten välisestä mahdollisesta tahallisesta tai tahattomasta vuorovaikutuksesta seuraa, että tietyt perusoletukset, joita yksinkertaisemmat satunnaistetut vertailukokeet edellyttävät, eivät välttämättä ole voimassa, kun halutaan arvioida yritystukien vaikuttavuutta. Yksi näistä tässä tilanteessa vaarantuvista perusoletuksista on SUTVA oletus: Tämä oletus tarkoittaa, että arvioitaessa tietyn toimenpiteen (tuen) vaikuttavuutta yksinkertaisella kahden ryhmän vertailuasetelmalla tarkastelun kohteena olevien yksiköiden välillä ei saisi olla suoraa tai epäsuoraa vuorovaikutusta. Mikäli tällaista vuorovaikutusta on, ei yksinkertaisilla satunnaistetuilla vertailukokeilla välttämättä saada estimoitua esimerkiksi toimenpiteen keskimääräistä vaikuttavuutta kuvaavaa ATE-parametria. Ongelmaksi voi muodostua esimerkiksi se, että verrattaessa toimenpiteen kohteeksi päätyneen ryhmän innovaatiotoiminnan tuloksellisuutta, koeryhmän innovaatiotoiminnan tuloksellisuudesta saadaan liian alhaisia arvioita, koska osa vertailuryhmään päätyneistä yrityksistä on hyötynyt innovaatiotukea saaneiden yritysten innovaatiotoiminnan tuloksena syntyvästä tiedosta (tiedon leviäminen). Samanaikaisesti tukea saavat yritykset voivat kehittää nopeammin uusia tuotteita, jonka myötä tukea ilman jäävien kilpailijoiden markkina-asema voi heikentyä ja niiden markkinaosuus pienentyä. Samantyyppisiä negatiivisia markkinavaikutuksia voi syntyä tuotanton tekijämarkkinoilla, jos esimerkiksi innovaatiotukea saava yritys saa tuen turvin palkattua parhaat asiantuntijat tai tutkijat.

Klusteroidut satunnaiskokeet (*clustered randomized experiments*) ovat yksi ratkaisu edellä kuvattuun arviointiongelmaan (ks. Athey ja Imbens 2017, ss. 109-114). Näissä satunnaiskokeissa toimenpidettä ei satunnaisteta havaintoyksikköjen tasolla, kuten esimerkiksi yrityksittäin, vaan kohdepopulaatio jaetaan ensin ryhmiin ("klustereihin"). Satunnaistaminen tapahtuu sitten ryhmätasolla siten, että kuhunkin ryhmään kuuluvat jäsenet saavat saman toimenpiteen eli päätyvät esimerkiksi toimenpiteen kohteeksi tai sitten ne jäävät kaikki vertailuryhmään. Mikäli ryhmät on muodostettu siten, että havaintoyksiköt voivat olla vuorovaikutuksessa vain ryhmien sisällä, voidaan toimenpiteiden kohteeksi valittujen ryhmien jäseniä verrata vertailuryhmien jäseniin suoraviivaisesti, tavanomaisia arviointimenetelmiä käyttäen. Tämä siis edellyttää, että vuorovaikutusta ei tapahdu eri ryhmiin kuuluvien yritysten välillä. Tällöin tarkastelua varten tehdään oletus osittaisesta (rajoitetusta) vuorovaikutuksesta (*partial interference*). Jos kiinnostuksen

kohteena olisi esimerkiksi yritysten vaikuttavuus hyvin pienissä yrityksissä, voisi yksi esimerkki satunnaistamisesta ryhmätasolla olla se, että yritykset jaetaan maantieteellisen sijaintinsa perusteella ryhmiin. Satunnaistaminen voisi tässä asetelmassa tapahtua siten, että tukia kohdennetaan alueittain hieman eri tavoin tai eri perusteilla ja että alueet olisi jaettu kohde- ja kontrolliryhmiin satunnaisesti.

3.1.2 Ulkoisvaikutusten mittaamisen haasteita

Toisaalta arvioitaessa innovaatiotukia ja niiden vaikuttavuutta kiinnostuksen kohde voi olla nimenomaan se, missä määrin toimenpiteestä aiheutuu ulkoisvaikutuksia (*externalities, spillovers*). Innovaatiotuen vaikuttavuutta arvioitaessa voi toisin sanoen olla oleellista tietää, miten innovaatiotukea saaneiden yritysten toimet ovat mahdollisesti vaikuttaneet myös sellaisten yritysten innovaatiotoimintaan, jotka eivät ole saaneet tukea. Tämän alueen tutkimus on melko uutta (ks esimerkiksi Duflo ja Saez 2003, , Athey and Imbens 2017, luku 11, Vazquez-Bare 2018).

Yksi erityishaaste ulkoisvaikutusten mittaamisessa satunnaistettujen koejärjestelyiden avulla on se, että se, mitä halutaan mitata, ei välttämättä ole selvää. Kyse on siis siitä, miten ulkoisvaikutus pitäisi määritellä arviointitilannetta varten ja missä määrin se on edes periaatteessa erotettavissa muista vaikutuksista.

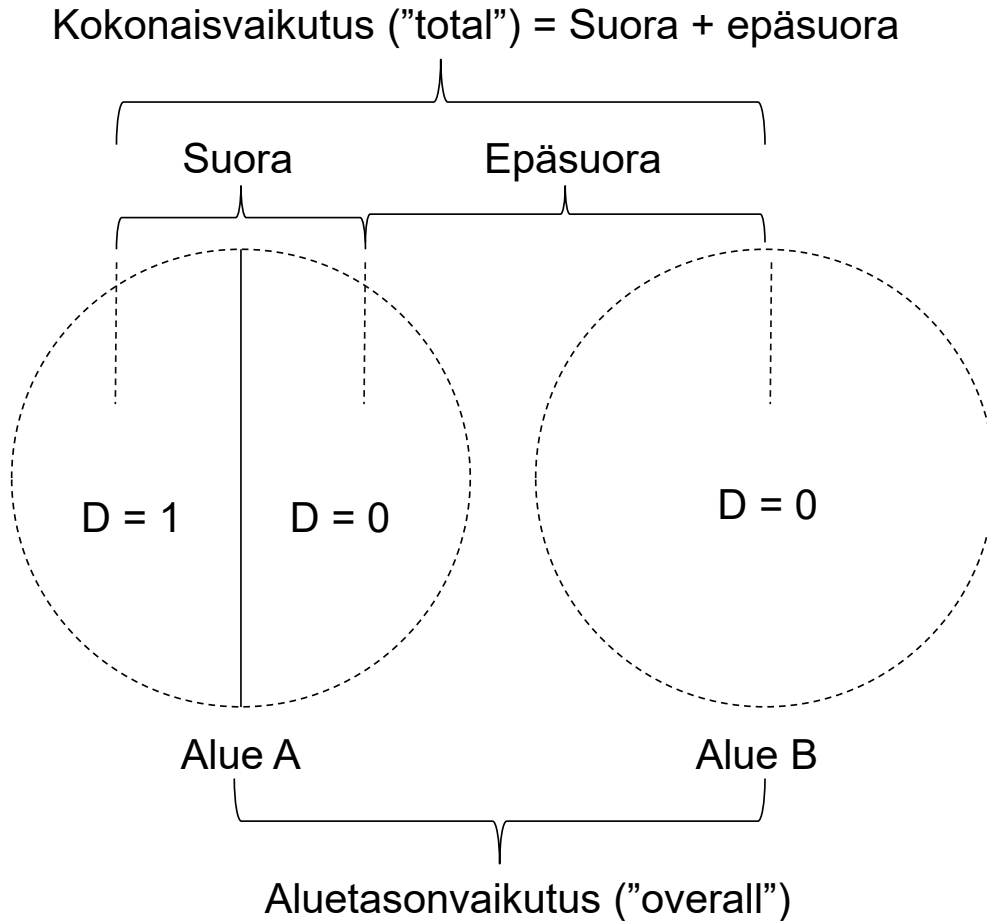
Kuvio 1 havainnollistaa erilaisia mahdollisia vaikutuksia silloin kun yritysten välillä voi olla vuorovaikutusta. Kuviossa on yksinkertaisuuden vuoksi esitetty kaksi aluetta, A ja B. Alueella A sijaitsevista yrityksistä satunnaisesti valittu osa saa julkista yritystukea ($D = 1$, eli kohderyhmä) ja osa ei ($D = 0$, vertailu- tai kontrolliryhmä). Alueella B sijaitsevista yrityksistä yksikään ei saa yritystukea. Oletus on, että alueiden sisällä voi olla yritysten välistä vuorovaikutusta, mutta että sitä ei tapahdu alueiden välillä. Ajatellaan lisäksi konkreettisuuden vuoksi, että tarkastelussa olevan yritystuen tarkoitus on lisätä työllisyyttä. Nyt voidaan erotella seuraavat vaikutukset (Halloran ja Hudgens 2016; ks. myös Vazquez-Bare 2018):

- **Kokonaisvaikutus** (*total effect*): Kun verrataan alueella A sijaitsevien yritysten työllistämistä kohderyhmässä ($D = 1$) alueella B sijaitsevien yritysten työllisyyteen (joista kukaan ei saa yritystukea), saadaan laskettua yritystuen keskimääräinen kokonaisvaikutus. Tämä vaikutus mittaa sitä, missä määrin alueen A yritystukea saaneet yritykset ovat lisänneet työllisyyttään verrattuna sellaisiin yrityksiin, jotka toimivat alueella, jolla kukaan ei ole saanut yritystukea. Tämän kokonaisvaikutuksen voidaan ajatella olevan ns. suorien ja epäsuorien vaikutusten summa. Nämä vaikutukset kuvataan alla.

- **Epäsuoravaikutus** (*indirect effect*): Kun verrataan alueella A sijaitsevien yritysten työllistämistä vertailuryhmässä ($D = 0$) alueella B sijaitsevien yritysten työllisyyteen (joista kukaan ei saa yritystukea), saadaan laskettua yritystuen keskimääräinen epäsuora vaikutus. Tämä vaikutus mittaa sitä, missä määrin alueen A yritystukea saaneet yritykset ovat vaikuttaneet saman alueen muihin yrityksiin, eli niihin jotka eivät ole saaneet yritystukea. Oleellista on, että tässä verrataan keskenään ei-tukea saaneita yrityksiä kahdessa erilaisessa ympäristössä: Toinen ryhmistä toimii ympäristössä, jossa tietty osuus yrityksistä on saanut yritystukea ja toinen ympäristössä, jossa kukaan ei ole saanut tukea.
- **Suora vaikutus** (*direct effect*): Kun vertaamme alueella A sijaitsevien yritysten työllistämistä kohderyhmässä ($D = 1$) ja vertailuryhmässä ($D = 0$), saadaan laskettua yritystuen keskimääräinen suora vaikutus. Vaikka vaikutuksen nimi saattaa olla hieman harhaanjohtava, tämä vaikutus mittaa sitä, missä määrin alueen A yritystukea saaneet yritykset ovat keskimäärin lisänneet työllisyyttään verrattuna saman alueen niihin yrityksiin, jotka eivät ole saaneet yritystukea. Oleellista on, että tässä verrataan keskenään tukea saaneita ja ei-saaneita yrityksiä siten, että molemmat ryhmät toimivat samanlaisessa ympäristössä (ts. ympäristössä, jossa tietty osuus yrityksistä on saanut yritystukea). Tässä on huomattava, että mikäli yritystukea saavat yritykset esimerkiksi syrjäyttävät (*crowding out*) työmarkkinoilla oman alueensa muita (ei-tukea-saaneita) yrityksiä, suora vaikutus on suurempi kuin mikäli tällaista syrjäyttämistä ei tapahtuisi.⁶
- **Aluetason (ryhmätason) yleisvaikutus** (*overall effect*): Kun verrataan kaikkien alueella A sijaitsevien yritysten työllistämistä (riippumatta siitä, ovatko kohde- vai vertailuryhmässä) alueella B sijaitsevien yritysten työllisyyteen, saadaan laskettua yritystuen keskimääräinen aluetason yleisvaikutus. Tämä vaikutus mittaa sitä, missä määrin alueen A yritykset ovat keskimäärin lisänneet työllisyyttään verrattuna alueen B yrityksiin. Tämä yleisvaikutus mittaa tietyllä tavalla yritystuen nettovaikutusta tietyn alueen (ryhmätason) näkökulmasta.

⁶ Halloranin ja Hudgensin (2016) tutkimus keskittyy pitkälti epidemiologian ja lääketieteen alan kysymyksiin. Heidän tutkimuksessaan kuvatussa esimerkissä toimenpide on ihmisten rokotus. Tässä yhteydessä rokotuksen prosentuaalinen kattavuus kohdejoukossa saa aikaan epäsuoria (välillisiä) vaikutuksia, koska rokottamattoman henkilön todennäköisyys törmätä sairastuneeseen pienenee, mitä suurempi osa muista ihmisistä saa rokotuksen. Tässä yhteydessä suora vaikutus vastaa seuraavaan kysymykseen: Mikä on rokotteen vaikutus, kun rokotteen kattavuus X%? Yritystukikontekstissa vastaava kysymys on: Mikä on tuen vaikutus, kun X% yrityksistä saa tukea?

Kuvio 1: Vaikuttavuusmekanismeja kun yritysten välillä on vuorovaikutusta



Lähde: Mukailleen Halloran ja Hudgens (2016, kuvio 1).

Kun arvioidaan yritystykä, kiinnostuksen kohteena on tyypillisesti edellä esitetyn va-
 lossa kokonaisvaikutus (total effect): Soveltavassa tutkimuksessa ja yritystykä käsitte-
 levässä kirjallisuudessa (ja politiikkakeskusteluissa) tätä vaikutusta kutsutaan kuitenkin
 usein suoraksi vaikutukseksi. Terminologia ei siis ole täysin vakiintunut.

Toinen erityishaaste ulkoisvaikutusten mittaamisessa satunnaistettujen koejärjestelyi-
 den avulla on se, että erilaisia mahdollisia ulkoisvaikutusten yhdistelmiä on hyvin mo-
 nenlaisia, koska ne saattavat riippua siitä, keiden ja minkäläisten yritysten kanssa jokin
 tietty yritys on vuorovaikutuksessa. Tätäkin haastetta on perusteltua pohtia hieman tar-
 kemmin.

Pohditaan esimerkkiä, jossa tarkastelun kohteena on tietty yksinkertainen vakiomuotoinen yritystuki. Tarkastelun kohteena voi olla vaikkapa tuen työllisyysvaikutus. Jotta voimme kuvata tarkasti, mitä vuorovaikutus voi tarkoittaa, hyödynnämme seuraavassa yksinkertaista matemaattista esitystapaa: Yritys voi joko saada tai olla saamatta tukea ja keskitytään kolmeen esimerkkiyritykseen. Voimme tässä kutsua näitä yrityksiä yrityksiksi 1, 2 ja 3. Merkitään lisäksi $Y_i(.,.,.)$:llä yrityksen i työllisyyttä: Tutkimuskirjallisuudessa termiä $Y_i(.,.,.)$ kutsutaan ”käsittelyvektoriksi” (*treatment vector*), ja sen avulla pidetään kirjaa siitä, mitä muille yrityksille voisi tapahtua silloin kun yritys i saa tai ei saa tukea. Kun siis pyritään periaatetasolla määrittelemään, miten yritystuki vaikuttaa jonkin yrityksen i työllisyyteen ($i \in \{1, 2, 3\}$), joudutaan mahdollisten vuorovaikutusten vuoksi miettimään, ketkä muut yritykset saavat tai eivät saa tukea.

Edellä sanotusta johtuen on kolmen yrityksen esimerkkitapauksessamme tarpeen pohdita seuraavia vaihtoehtoisia mahdollisuuksia: $Y_i(1,1,1)$, $Y_i(1,1,0)$, $Y_i(1,0,1)$, $Y_i(1,0,0)$, $Y_i(0,1,1)$, $Y_i(0,1,0)$, $Y_i(0,0,1)$, $Y_i(0,0,0)$. Näistä ensimmäinen eli $Y_i(1,1,1)$ kuvaa esimerkiksi yrityksen i työllisyyttä tilanteessa, jossa kaikki yritykset saavat yritystukea, $Y_i(1,0,0)$ kuvaa tilannetta, kun vain yritys i saa yritystukea, mutta kukaan muu ei saa, ja $Y_i(0,0,0)$ puolestaan kuvaa yrityksen i työllisyyttä tilanteessa, jossa yksikään kolmesta yrityksestä ei saa yritystukea. Muut termit voidaan tulkita vastaavasti, samaa ajattelutapaa noudattaen.

Edellä esiteltyä merkintätapaa hyödyntämällä voidaan sanoa, että erotus $Y_i(1,0,0) - Y_i(0,0,0)$ kuvaa yritystuen kausaalivaikutusta yrityksen i työllisyyteen (ts. sen kasvua tai vähenemistä), kun vain yritys i saa yritystukea. Mikäli yrityksiä välillä on vuorovaikutusta ja mikäli yrityksen i tuesta saama hyöty on vähäisempi, jos muut yritykset saavat samaa tukea, voisi tällöin esimerkiksi vallita tilanne: $Y_i(1,0,0) - Y_i(0,0,0) > Y_i(1,1,0) - Y_i(0,1,0) = Y_i(1,0,1) - Y_i(0,0,1) > Y_i(1,1,1) - Y_i(0,1,1)$. Näistä ensimmäinen epäyhtälö kertoo, että verrattuna tilanteeseen, jossa vain yritys i saa tukea, yrityksen i tuesta saama hyöty on alhaisempi, jos yritys j tai yritys k saa myös tukea ($j \neq k \neq i$). Jälkimmäinen yllä olevista epäyhtälöistä puolestaan kertoo, että yrityksen i tuesta saama hyöty voi olla vieläkin alhaisempi, kun yritykset j ja k molemmat saavat tukea.⁷ Tämä voisi olla esimerkiksi seurausta siitä, että kun myös muut yritykset saavat yritystukea, yritykset päätyvät todennäköisemmin kilpailemaan työntekijöistä työmarkkinoilla. Myös muita vastaavia vertailuja voidaan tehdä.

⁷ Olkoon $i = 1$, $j = 2$ ja $k = 3$. Nyt esimerkiksi epäyhtälö $Y_1(1,0,0) - Y_1(0,0,0) > Y_1(1,1,0) - Y_1(0,1,0)$ kuvaa sitä, että verrattuna tilanteeseen, jossa vain yritys 1 saa tukea, yrityksen 1 tuesta saama hyöty on alhaisempi, jos yritys 2 saa myös tukea. Toisaalta yhtälö $Y_1(1,1,0) - Y_1(0,1,0) = Y_1(1,0,1) - Y_1(0,0,1)$ kuvaa sitä, että kun yritys 1 saa tukea, yrityksen 1 tuesta saama hyöty ei riipu siitä, kumpi yrityksestä 2 tai 3 saa myös tukea. Hyöty on siis yhtä suuri silloin, kun yritys 2 saa myös tukea (mutta yritys 3 ei) kuin silloin, kun yritys 3 saa myös tukea (mutta yritys 2 ei).

Yllä esitelty merkintätapa on hyödyllinen siksi, että sen avulla nähdään, että jos vuorovaikutuksessa olevia yrityksiä on enemmän kuin kolme, kasvaa erilaisten mahdollisten tulemien määrä helposti hyvin suureksi: Mikäli tarkasteluun tulevia yrityksiä on N kappaletta, erilaisten mahdollisten käsittelyvektoreiden määrä on 2^N . Edellä $N = 3$, joten erilaisia käsittelyvektoreita oli esimerkiksi kahdeksan kappaletta ($2^3 = 8$). Koska tämä luku kasvaa nopeasti, kun N kasvaa, ulkoisvaikutusten mittaamista varten täytyy usein tehdä yksinkertaistavia oletuksia. Tämä on tärkeä asia, joka on ymmärrettävä ja otettava huomioon, kun koeasetelmia suunnitellaan käytännössä.

Osittaisen vuorovaikutuksen (*partial interference*) oletus yhdistettynä vaihdettavuusoletukseen (*exchangeability assumption*) on yksi tapa ratkaista mittaamisongelmat, joita yritysten välisestä vuorovaikutuksesta syntyy (Halloran ja Hudgens 2016, Vasquez-Bare 2018). Nämä oletukset tarkoittavat, että kun pyritään periaatetasolla määrittelemään, miten yritystuki vaikuttaa jonkin yrityksen i työllisyyteen, ei jouduta miettimään, mitkä tietyt muut yritykset saavat tai eivät saa tukea, vaan mikä osuus yrityksen i alueen (ryhmän) yrityksistä saa tukea. Kun tämä pätee, esimerkiksi seuraava havainnollistava kaksivaiheinen kontrolloitu satunnaiskoe saattaisi olla tapa mitata epäsuoria innovaatiotukien vaikutuksia (Halloran ja Hudgens 2016, Athey ja Imbens 2017, s. 133):

- Kuvitellaan, että yritykset voidaan jakaa (maantieteellisen sijaintinsa tai muun tekijän perusteella) esimerkiksi G määrään ryhmiä siten, että ryhmien välillä ei ole tai on mahdollisimman vähän vuorovaikutusta.
- Ensimmäisessä vaiheessa puolet ryhmistä valitaan satunnaisesti siten, että niihin kohdistuu toimenpide eli että tietylle osuudelle ryhmän jäseniä kohdistetaan yritystukea. Loput ryhmistä ovat sellaisia, että yksikään näihin ryhmiin kuuluva yritys ei saa yritystukea. Toisessa vaiheessa $X\%$ kunkin toimenpiteen kohteeksi valitun ryhmän yrityksistä saa yritystukea ja $(1-X)\%$ ei saa sitä. Myös tämä tuen kohdentaminen yritystasolla tehdään satunnaisesti.
- Tuen suora, epäsuora ja kokonaisvaikutus voidaan nyt arvioida tavalla, joka esiteltiin edellä Kuvion 1 yhteydessä. Esimerkiksi tuen epäsuoravaikutus saadaan arvioitua, kun verrataan tukea saaneilla alueilla sijaitsevien mutta vertailuryhmään kuuluvien yritysten tulemia niillä alueilla sijaitsevien yritysten tulemiin, joilla kukaan ei saa yritystukea. Tämä vaikutus mittaa sitä, missä määrin yritystuki on vaikuttanut niihin yrityksiin, jotka eivät ole saaneet yritystukea mutta jotka sijaitsevat alueella, joissa yritystukea saaneita yrityksiä on.

Ulkoisvaikutuksiin melko läheisesti liittyvä mutta kuitenkin erillinen ilmiö on ns. ryhmävaikutukset (*peer effects*). Yritystukiin sovellettuna ryhmävaikutus voisi liittyä esimerkiksi siihen, että tietty yksittäinen yritys kokee tarpeelliseksi hakea innovaatiotukea siksi, että sen toimintaympäristössä muut yritykset ovat käynnistämässä innovaatiotoimintaa ja hakeneet innovaatiotukea. Tämänkaltaisten ryhmävaikutuksien mittaaminen ja identifiointi tilastoaineistoista on osoittautunut visaiseksi ongelmaksi (ks. Manski

1993, Angrist 2014, Sacerdote 2014). Yksi haaste liittyy siihen, että vuorovaikutuksen merkityksen arvioinnissa on usein kyse siitä, miten tietyn yksittäisen yrityksen käyttäytymiseen vaikuttaa sen ryhmän käytös, johon ko. yritys itsekin kuuluu. Kyse on tietynlaisesta peilikuvaongelmasta (*reflection problem*): Kun tarkasteluun otettavien havaintoyksiköiden välillä on vuorovaikutusta, on vaikea erottaa, vaikuttaako joukon keskimääräinen käyttäytyminen siihen kuuluvien yksilöiden käyttäytymiseen vai heijasteleeko sekä joukon keskimääräinen käyttäytyminen että siihen kuuluvien yksiköiden käyttäytyminen jotain niiden yhteistä piirrettä tai ominaisuutta (*exogenous effects*) tai jotain muuta niille kaikille yhteistä ulkoista taustatekijää (*contextual effects*). Tällaisia piirteitä tai taustatekijöitä voi olla vaikea mitata tai havaita aineistoista, joita tutkijoilla on käytettävissä.

3.2 Erityiskysymys 2: Koejärjestelyn toteuttaminen, kun toimenpiteen kohteeksi joutuvan omilla toimilla on merkitystä sille, miten toimenpide vaikuttaa

Yritystukien vaikuttavuuden arviointi voi monimutkaistaa myös se, että koejärjestelyn vaikutusten mittaamiseen vaikuttaa se, miten paljon toimenpiteen kohteeksi valittu näkee vaivaa sen eteen, että toimenpide tulee toteutetuksi oikealla tavalla. Jos esimerkiksi tietynlaisista innovaatiotukea saanut yritys arvelee ja odottaa, että tuesta on sille vain vähän hyötyä, se ei välttämättä koe perustelluksi kohdentaa voimavaroja, kuten esimerkiksi osaavien työntekijöiden työaika, saadun tuen hyödyntämiseksi tehokkaasti. Tämä mahdollisuus on perusteltua ottaa huomioon, kun yritystukiin liittyviä koeasetelmia suunnitellaan käytännössä: Jos siis tietyn tuen vaikuttavuutta mitataan yksinkertaisella satunnaistetulla koejärjestelyllä ja jos tulos kertoo, että tuen vaikuttavuus on vähäinen, se voi johtua joko siitä, että tuki on aidosti tehoton tai siitä, että tukea saaneet yritykset odottavat ja arvioivat tuen hyödyn vähäpätöiseksi, eivätkä siksi panosta sen hyödyntämiseen (Chassang, Miguel ja Snowberg 2012, Hanna, Duflo ja Greenstone 2016).

Tutkimuskirjallisuudessa on pohdittu, kuinka toimenpiteen (tuen) vaikuttavuutta voitaisiin arvioida satunnaistettujen kontrolloitujen koejärjestelyiden avulla yllä kuvatussa tilanteessa. Chassang, Miguel ja Snowberg (2012) kutsuvat kehittelemäänsä tutkimusasetelmaa valikoivaksi koejärjestelyksi (*selective trial*). Valikoivan koejärjestelyn tavoite on parantaa koejärjestelyn tulosten yleistettävyyttä (ulkoista validiteettia) vaarantamatta sen kykyä mitata tarkasti haluttua asiaa (sisäistä validiteettia). Ajatus on, että

koejärjestelyn toteuttaa taho ("päämies"), jonka tavoite on saada mahdollisimman paljon tietoa toimenpiteen vaikuttavuudesta tilanteessa, jossa toimenpiteen kohteeksi (mahdollisesti) valittavilla on virheellisiä näkemyksiä toimenpiteen tehokkuudesta (hyödyllisyydestä) ja/tai taipumus olla panostamatta toimenpiteen toteutukseen.

Valikoivassa koejärjestelyssä yrityksille annetaan mahdollisuus ilmaista näkemyksensä toimenpiteen hyödyllisyydestä siten, että ne tämän ilmaistun hyödyllisyyden myötä valikoituvat todennäköisemmin päätymään toimenpiteen kohteeksi. Tämä näkemyksen ilmaisu edellyttää yritykseltä voimavarojen käyttöä siten, että ne yritykset, jotka kokevat toimenpiteen hyödyllisemmäksi, ovat valmiit "maksamaan" toimenpiteestä enemmän. Käytännössä tämä voisi tarkoittaa, että esimerkiksi yritykset, jotka kokevat tuen hyödylliseksi, maksavat päästäkseen toimenpiteen kohteeksi. Tämä voi tarkoittaa rahamääräistä maksua (eli käytännössä yrityksen näkökulmasta alennettua tukimäärää) tai muuta ei-rahamääräistä panostusta, kuten esimerkiksi työvoiman käyttöä hakemuksen täyttämiseen. Valikoiva koejärjestely toimii siten, että yritykset, jotka ilmaisevat halunsa maksaa enemmän, saavat tuen suuremmalla todennäköisyydellä.

Valikoivassa koejärjestely saattaisi esimerkiksi soveltua innovaatioasetelin vaikuttavuuden arviointiin. Innovaatioaseteli voidaan kohdentaa pk-yrityksille, joilla on "uusi kansainvälistä kasvupotentiaalia omaava tuote- tai palveluidea". Julkisen sektorin tukema innovaatioaseteli on tarjolla siksi, että ilman ulkopuolista osaamista tätä potentiaalia ei välttämättä onnistuta hyödyntämään. Yksi mahdollinen syy innovaatioasetelin vaatimattomaan vaikuttavuuteen voisi olla se, että se kohdentuu yrityksille, jotka eivät arvioi saavansa siitä erityisen suurta hyötyä. Ne kyllä ottavat julkisesti tuetun setelin vastaan, mutta eivät välttämättä panosta omia voimavarojaan kun ne käyttävät seteliä: Mikäli näin käy, setelillä ostettavasta palvelusta ei saada suurinta mahdollista hyötyä.

Yksi tapa toteuttaa edellä mainittu valikoiva koejärjestely jaettaessa esimerkiksi 5000€ innovaatioaseteliä yrityksille on se, että sen kohdentamisessa hyödynnetään ns. Becker-DeGroot-Marschak (BDM) mekanismia (Berry, Fischer ja Guiteras 2018).

BDM mekanismi voisi toimia esimerkiksi seuraavasti: Ennen koetta selvitetään suurin mahdollinen arvio tuen (tässä: innovaatioasetelin) hyödyllisyydestä yrityksille. Tätä euromääräistä arvoa merkitään alla X:llä ja se täytyy määritellä ennen kuin koe tehdään.⁸

⁸ X:n arvo on selvitettävä tapauskohtaisesti etukäteen käytettävissä olevan tiedon ja ymmärryksen valossa. Kokeen onnistumisen kannalta oleellista ei ole tarkka X:n arvo vaan se, että se on jonkin verran korkeampi kuin mitä yritykset voisivat olla halukkaita innovaatioasetelistä tarjoamaan. On hyvä huomata, että X€ voi myös olla suurempi kuin 5000€, jos arvioidaan, että innovaatioasetelillä saa hankittua palveluita, joiden hyöty ainakin joillekin yrityksille on suurempi kuin 5000€.

Kun X on määritelty, yrityksille kuvataan heidän kohtaamansa valintatilanne huolella. Yritysten kohtaama valintatilanne sisältää seuraavat vaiheet:

- Vaihe 1: Innovaatioseteliä hakeneelta yritykseltä kysytään, kuinka paljon se on valmis maksamaan saadakseen sen. Merkitään yrityksen antamaa tarjousta b :llä.
- Vaihe 2: Tämän jälkeen setelille arvotaan hinta, joka tässä tapauksessa olisi satunnaisluku väliltä $[0\text{€}, X\text{€}]$, jossa $X\text{€}$ on edellä mainittu suurin mahdollinen arvio tuen hyödyllisyydestä yrityksille ja joka ei siis riipu siitä, mitä yritykset tässä vaiheessa tarjoavat. Merkitään arvottua hintaa x :llä.
- Vaihe 3: Mikäli arvottu hinta on yhtä suuri tai alhaisempi kuin yrityksen ilmoittama maksuhalukkuus (*willingness-to-pay*, WTP) eli mikäli $b \geq x$, yritys saa ”ostaa” innovaatiosetelin arvottuun hintaan. Tämä arvottu hinta on siis x . Jos arvottu hinta on korkeampi kuin yrityksen ilmoittama maksuhalukkuus eli mikäli $b < x$, yritys ei saa innovaatioseteliä. On tärkeää huomata, että yrityksen ilmoittama maksuhalukkuus vaikuttaa tässä vain todennäköisyyteen, että saako se innovaatiosetelin (eli todennäköisyyteen että $b \geq x$), mutta ei hintaan, jonka se siitä maksaa (x). Yritys ei siis vaikuta innovaatiosetelin hintaan tehdessään tarjouksen vaiheessa 2.

BDM toimii toisin sanoen siten, että yritykset, jotka ilmaisevat halunsa maksaa enemmän, saavat tuen suuremmalla todennäköisyydellä. Yritykset tietävät tämän ja ymmärtävät, että suurempi maksuhalukkuus lisää todennäköisyyttä, että saa yritystuen. BDM mekanismin ominaisuus on myös se, että se kannustaa (rationaalisesti toimivaa) yritystä ilmoittamaan maksuhalukkuutensa rehellisesti (*incentive compatible*).

Kuten Berry, Fischer ja Guiteras (2018) kuvaavat, edellä kuvatun asetelman etuja ovat mm. seuraavat:

- *Maksuhalukkuus ja kysyntä*: Asetelma tuottaa tietoa yritysten maksuhalukkuudesta (WTP-määrästä) eli kvantitatiivista tietoa siitä, minkä verran innovaatioseteleihin kohdistuu kysyntää pk-yrityksien keskuudessa.
- *Vaikutavuus*: Asetelma tekee mahdolliseksi innovaatiosetelin kausaalivaikutuksen estimoinnin: sen avulla voidaan laskea innovaatiosetelin keskimääräinen vaikuttavuus esimerkiksi yrityksen (jollain tavoin mitatun) innovaatiotoiminnan laajuuteen tai sen tuloksellisuuteen sekä arvioida tähän vaikutukseen sisältyvää vaihtelua ja heterogeenisyyttä ns. marginaalivaikutuslaskelmien (*marginal treatment effects*) avulla.
- *Innovaatio- ja kasvatulemat*: Asetelman avulla voidaan myös arvioida, missä määrin yrityksillä on tarvetta (kysyntää) sellaisille innovaatioihin ja kansainväliseen kasvuun liittyville tuleville, joita innovaatioseteleillä pyritään saamaan aikaan. Hieman yksinkertaistaen voidaan sanoa, että asetelman avulla voidaan laskea, kuinka paljon pk-yritykset ovat valmiit maksamaan saadakseen lisää kansainvälistä liikevaihtoa tai patentin.

3.3 Erityiskysymys 3: Syyseuraus-suhteiden mekanismien ymmärrys

Satunnaistettujen kontrolloitujen koejärjestelyiden avulla halutaan usein selvittää, mikä jonkin toimenpiteen vaikutuksen keskimääräinen vaikutus on. Kiinnostuksen kohteena on toisin sanoen se, onko vaikutusta lainkaan ja jos on, mikä sen koko (suuruus) on. Tavoitteena ei useinkaan ole sen selvittäminen tarkemmin, mistä syy-seuraussuhteista eli kausaalista vaikutuskanavasta (*causal mechanism*) vaikutus syntyy. Arvioitaessa yritystukien vaikutusta on luontevaa kysyä myös sitä, miksi ne vaikuttavat siten kuin vaikuttavat, mikäli joitain vaikutuksia ylipäättään on. Kyse on syy-seurausprosesseista (*causal process*) joiden myötä yritystuki saa aikaan tietyn vaikutuksen. Kiinnostuksen kohteena on siis välitysmekanismista (*causal mediation effect*), joka määrittelee, missä määrin toimenpide vaikuttaa tulemiin tietyn välittävän muuttujan (*mediator*) kautta.

Uusin tilastollinen ja ekonometrinen tutkimus on arvioinut sitä, miten satunnaistetut vertailukokeet soveltuvat vaikutusmekanismien selvittämiseen. Tähän liittyen mm. Imai, Tingley ja Yamamoto (2013) ovat korostaneet sitä, jos tavanomaisten satunnaistettujen kontrolloitujen koejärjestelyiden (*single-experiment design*) avulla halutaan selvittää vaikutuksen taustalla olevia syy-seuraussuhteita, joudutaan tekemään useita oletuksia, joiden pitävyyttä ei voida tilastoaineiston avulla arvioida. Tämä on tärkeää ymmärtää, kun koeasetelmia suunnitellaan käytännössä ja kun kiinnostuksen kohteena on tukien vaikutusmekanismit.

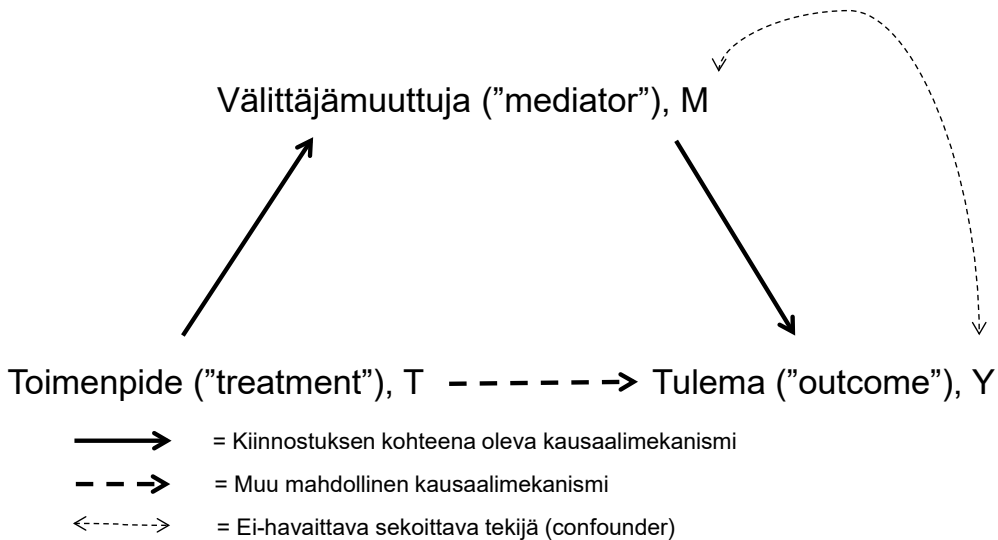
Kuvio 2 havainnollistaa erilaisia mahdollisia vaikutusmekanismeja silloin kun yritysten välillä ei ole vuorovaikutusta. Kuviossa on esitetty tilanne, jossa toimenpiteen (T) kausaalivaikutus kiinnostuksen kohteena olevaan tulemaan (Y) välittyy välittäjämuuttujan (M) kautta. On mahdollista, että toimenpide vaikuttaa tulemaan myös muita kanavia pitkin: Niitä kuvataan katkoviivanuolella. Toimenpiteen kokonaisvaikutuksen voidaan siten ajatella koostuvan epäsuorasta vaikutuksesta (*indirect effect*), joka etenee välittäjämuuttujan kautta ja jonka olemassa oloa tässä halutaan selvittää, ja suorista muista vaikutuksista (*direct effect*), joka syntyy muista mahdollisista syy-seuraussuhteista, jotka sitovat toimenpiteen tulemaan. Kirjallisuudessa terminologia on hieman vakiintumaton ja osin päällekkäinen, joten on syytä korostaa tässä, että tässä kuvatut suorat ja epäsuorat vaikutukset eivät ole samoja kuin ne, jotka aiheutuvat yritysten välisestä vuorovaikutuksesta. Usein on tarpeen olettaa, että yritysten välistä vuorovaikutusta ei ole, jotta voidaan tunnistaa kausaalimekanismeja.

Tarkastellaan konkreettisuuden vuoksi tilannetta, jossa kiinnostuksen kohteena on jonkin tietyn työllistämistuen eli esimerkiksi ns. palkkatuen (T) vaikutus työllisyyteen (Y) ja jossa mahdollinen välittäjämuuttuja on yhden uuden työntekijän palkkaukseen ja rekrytointiin liittyvät kokonaiskustannukset (M). Yritystasolla ajateltuna epäsuora ja suora vaikutus vastaavat seuraaviin kontrafaktuaalisiin kysymyksiin:

- **Epäsuora vaikutus:** Miten yrityksen työllisyys muuttuisi, jos työllistämisen kokonaiskustannuksia voitaisiin muuttaa tuetulta tasolta tavalliselle (ei-tuetulle) tasolle, *pitäen työllisyystuki muuttumattomana?*
- **Suora vaikutus:** Miten yrityksen työllisyys muuttuisi, jos yritys saisi työllistämistukea verrattuna tilanteeseen, jossa se ei sitä saa, *pitäen työllistämisen ja rekrytoinnin kokonaiskustannukset muuttumattomina?*

Jos työllisyydellä ei ole minkäänlaista muuta (suoraa) vaikutusta, tuen koko vaikutus syntyy siitä, että se alentaa rekrytoinnin ja työllistämisen kustannuksia. Jos työllistämistukea saava yritys kuitenkin kiinnittää tuen saannin vuoksi huomiota työllistämiseensä muistakin syistä kuin kustannusvaikutusten vuoksi, työllistämistuella on – Kuviossa 2 määritellyllä tavalla – suoria vaikutuksia.

Kuvio 2: Suoria ja epäsuoria vaikuttavuusmekanismeja



Lähde: Mukailen Ima, Tingley ja Yamamoto (2013, kuvio 1).

Kirjallisuudessa on esitelty jonkin verran erilaisia räätälöityjä satunnaistettuja koejärjestelyitä, joiden avulla voidaan arvioida toimenpiteiden vaikutusten taustalla olevia syyseuraussuhteita ja -mekanismeja. Osa tutkimuksista, joissa näitä räätälöityjä koejärjestelyitä kuvataan, on teknisesti haastavia ja siksi melko vaikeasti ymmärrettäviä.

Tästä huolimatta nämä tutkimukset ovat tärkeitä, sillä ne tarjoavat kehikon, joiden avulla käytännön koejärjestelyitä voidaan suunnitella.

Esittelemme esimerkkinä asetelman, jota Ima, Tingley ja Yamamoto (2013) kutsuvat yhtäaikaisen rohkaisun asetelmaksi (*parallel encouragement design*). Kiinnostuksen kohteena on keskimääräinen epäsuoravaikutus tunnollisille yrityksille (*average indirect effect for compliers*; ks. myös Angrist, Imbens ja Rubin 1996). Tunnollisuus viittaa tässä sellaisiin yrityksiin, jotka reagoivat asetelmaan sisältyvään rohkaisuun odotetulla ja tarkoitetulla tavalla. Kuvaamme alla tarkemmin, mitä tämä tarkoittaa.

Konkreettisuuden vuoksi sovellamme yhtäaikaisen rohkaisun asetelmaa seuraavassa palkkatukeen. Koska palkkatukea saava yritys joutuu noudattamaan palkkatukipäätöksen ehtoja, voi esimerkiksi jo pelkästään sillä, että se joutuu tutustumaan näihin ehtoihin osana palkkatuen hakuprosessia, olla vaikutus yrityksen rekrytointipäätöksiin. Tämä ei ole ainoastaan teoreettinen mahdollisuus, sillä aikaisemmissa tutkimuksissa on havaittu, että jo pelkästään kyselyyn vastaaminen saattaa muuttaa ihmisten käyttäytymistä verrattuna tilanteeseen, jossa he eivät vastaisi kyselyyn.

Kuviota 2 mukaillen voidaan nyt todeta, että kiinnostuksen kohteena on se, että aiheutuuko palkkatuen (T) vaikutus työllisyyteen (Y) siitä, että se alentaa uuden työntekijän palkkaukseen ja rekrytointiin liittyviä kokonaiskustannuksia (M). Palkkatuen vaikuttavuuden ymmärtäminen edellyttää näin ollen sitä, että kyetään erikseen arvioimaan sen suora ja epäsuora vaikutus:

- Epäsuora vaikutus on se, josta ollaan kiinnostuneita: Tässä ko. vaikutus olisi se, missä määrin palkkatuki vaikuttaa yritysten työllisyyteen *sen ja vain sen vuoksi*, että se alentaa palkkauksen ja rekrytointin kustannuksia.
- Suora vaikutus liittyy muihin mahdollisiin palkkatuen kausaalivaikutuksiin: Jos sillä, että yritys joutuu tutustumaan palkkatukipäätöksen ehtoihin osana palkkatuen hakuprosessia, on vaikutus yrityksen rekrytointipäätöksiin, olisi kyseessä palkkatuen muu kausaalivaikutus. Tämä on luonnollisesti vain yksi esimerkki mahdollisista muista kausaalivaikutuksista.

Yhtäaikaisen rohkaisun asetelma koostuu kaksiosaisesta satunnaistamisesta ja se toimii seuraavasti:

- Osa 1: Yritykset jaetaan satunnaisesti kahteen ryhmään, joita voidaan tässä kutsua ryhmiksi A ja B.
- Osa 2: Ryhmät A ja B saavat erilaisen käsittelyn:
 - Osalle ryhmään A kuuluvista yrityksistä kohdennetaan palkkatukea ja lopuille ei. Ne, jotka ryhmässä A saavat palkkatukea ja jotka eivät, valitaan satunnaisesti.

- Ryhmään B kuuluvat yritykset saavat joko palkkatukea tai vaihtoehtoisesti niitä ”rohkaistaan” (jollain perustellulla tavalla) alentamaan rekrytointiin ja palkkaamiseen liittyviä kokonaiskustannuksia. Ne, jotka ryhmässä B saavat palkkatukea ja keitä rohkaistaan, valitaan satunnaisesti. Tässä ”rohkaus” voi esimerkiksi tarkoittaa yleisen, rekrytointia helpottavan tiedon tarjoamista.⁹

Yhtäaikaisen rohkaisun asetelma nojaa kahteen oletukseen: Ensinnäkin, oletus on, että mikäli yritykset reagoivat kannustavaan rohkaisuun, ne aidosti pyrkivät alentamaan rekrytointikustannuksiaan. Oletus siis on, että yritykset eivät reagoi kannustukseen epäintuitiivisesti tekemällä juuri päinvastoin eli tekemällä jotain, joka nostaa rekrytointikustannuksia. Tätä käyttäytymisoletusta kutsutaan monotonisuusehdoksi (*monotonicity*). Tunnolliset yritykset (*compliers*) ovat yrityksiä, jotka reagoivat rohkaisuun tekemällä valintoja, jotka alentavat rekrytointikustannuksia. Toiseksi, tärkeää asetelmassa on se, että rohkaisu ei mitenkään suoraan liity niiden ihmisten palkkaamiseen, joita palkkatuella voisi olla tarkoitus palkata. Tätä ehtoa kutsutaan poissulkemisehdoksi (*exclusion restriction*). Tämä tieto voisi mm. olla sellaista, että se alentaa rekrytointikustannuksia vähentämällä tyypilliseen rekrytointiin tarvittavien henkilötyöntuntien määrää.

Kuten yllä todettiin, epäsuora vaikutus kuvaa tässä yhteydessä kiinnostuksen kohteena olevaa syy-seuraussuhdetta, joka selittää osaltaan, miksi palkkatuella mahdollisesti on vaikutus yritysten työllisyyteen. Ko. vaikutus viittaa yllä kuvatussa asetelmassa siihen, miten yllä mainittujen tunnollisten yritysten työllisyys keskimäärin muuttuu siksi, että palkkatuki aiheuttaa rekrytoinnin ja työllistämisen kokonaiskustannuksiin vaihtelua. Tämä vaikutus ei sisällä sitä, että palkkatuen saaminen saattaa muuttaa yritysten työllisyyttä myös muista kuin rahamääräisistä syistä.

Ima, Tingley ja Yamamoto (2013, s. 20) määrittelevät tutkimuksessaan tarkemmin suureen, joka edellä kuvatun yhtäaikaisen rohkaisun asetelman avulla voidaan identifioida. Käytännössä tämä tarkoittaa, että he kuvaavat tarkemmin matemaattisen kaavan, jonka avulla palkkatuen epäsuora vaikutus voitaisiin laskea. Koska ko. kaava on monimutkainen ja koska sen esittely vaatisi uudenlaisen matemaattisen notaation hyödyntämistä, kaavaa ei esitetä tässä. Kaava perustuu havaittuihin eroihin työllistämässä erilaisen käsittelyn saaneiden yritysten välillä ryhmissä A ja B.

⁹ Rohkaus voi tarkoittaa eri asioita eri tilanteissa ja riippuu tuki-instrumentista, jonka vaikuttavuutta halutaan selvittää.

3.4 Muita erityiskysymyksiä

Kun halutaan selvittää yritystukien vaikuttavuutta, satunnaistettujen koejärjestelyiden avulla, voi olla tarpeen pohtia tiettyjä muitakin erityiskysymyksiä. Luettelemme niistä muutamia mielestämme tärkeimpiä tässä:

- *Skaalautuvuus*: On tilanteita, joissa joitain tiettyä toimenpidettä ja vaikuttavuusarviointimenettelyä kokeillaan ensin pienimuotoisesti siten, että pilotti-projektissa kohde- ja vertailuryhmä muodostavat vain pienen osan toimenpiteen mahdollisesta kohdepopulaatiosta. Banerjee ym. (2016) korostavat, että pienimuotoisesti toteutettua satunnaistettua koejärjestelyä ei aina välttämättä voida toteuttaa yhtä tehokkaasti tai samalla tavoin suuremmassa mittakaavassa (ts. se ei skaalaudu). Kyse on siis siitä, missä määrin satunnaistetulla koejärjestelyllä saatavat tulokset tarkastelussa olevan tuen vaikuttavuudesta saattavat muuttua, kun toimenpidettä laajennetaan ja kun se ulotetaan kattamaan suurempi yritysjoukko. Skaalautuvuus riippuu sekä siitä, kuinka paljon yritysten välille syntyy vuorovaikutusta mittakaavan kasvaessa (mm. ulkoisvaikutusten vuoksi, tai markkinoilla syntyvien tasapainovaikutusten vuoksi) ja siitä, missä määrin toimenpide ja koejärjestely onnistutaan toteuttamaan yhtä huolellisesti suuressa mittakaavassa kuin miten se voidaan tehdä pienimuotoisemmassa pilottiprojektissa (ks. myös Banerjee ja muut 2017, s. 166-167).
- *Kontekstisidonnaisuus*: On tunnettua, että toimenpiteen vaikuttavuus voi vaihdella yrityksestä toiseen (*heterogeneous treatment effects*), kun vaikuttavuutta tarkastellaan tietyssä ympäristössä ja tilanteessa. Vivald (2015) korostaa, että toimenpiteiden vaikuttavuus saattaa myös olla vahvasti tilannesidonnaista (*context-dependence of impact evaluation*). Kyse on tulosten yleistettävyydestä ja siitä, missä määrin jokin tuen vaikuttavuutta arvioiva satunnaistettu tutkimus tuottaa saman tuloksen kuin jokin aiempi, saman toimenpiteen vaikuttavuutta hieman toisenlaisessa yhteydessä selvittänyt tutkimus. Mikäli toimenpiteiden vaikuttavuus on vahvasti tilannesidonnaista, ei tietyn tuen vaikuttavuudesta tietyssä tilanteessa voida välttämättä päätellä, miten se vaikuttaa toisenlaisessa toimintaympäristössä (ks. myös Deaton 2010).

- *Erityiset tutkimuskysymykset:* Yritystukijärjestelmän kehittämiseen ja ohjaamiseen liittyy myös tutkimuskysymyksiä, joihin ei voida aina suoraviivaisesti vastata satunnaistettujen koejärjestelyiden avulla. Satunnaistetut koasetelmat soveltuvat erityisen hyvin keskimääräisten yritystason kokonaisvaikutusten vaikutusten arviointiin. Niiden avulla on kuitenkin vaikeampaa ja usein jopa mahdotonta arvoida esimerkiksi sitä, kuinka jotkin politiikan kannalta oleelliset taustatekijät vaikuttavat yrityksen päätökseen hakea ohjelmaan. Esimerkiksi yrittäjän perhetaustaan tai yleiseen taloustilanteeseen arviointikoe ei voi vaikuttaa, jolloin niiden merkitystä kokeeseen osallistumiselle ei voida kokeiden avulla luotettavasti arvioida. (ks. mm. Heckman ja Smith 1995). Toinen esimerkki kysymyksestä, johon arviointikoe ei anna täydellistä vastausta, on kokonaistaloudellisten vaikutusten arviointi. Koska kokonaisia talouksia ei luonnollisestikaan voida jakaa satunnaisesti koe- ja verrokkiryhmiin, tämän tason vaikutusten arviointi on pois suljettu. Yritystukien suorien ja epäsuorien vaikutusten arvoinnilla voidaan silti tuottaa myös kokonaistaloudellisen arvion kannalta tärkeitä tuloksia mm. vaikutusmekanismeista.

4 Satunnaistettu vaikuttavuusarviointi yritystukipolitiikan kehittämistyökaluna

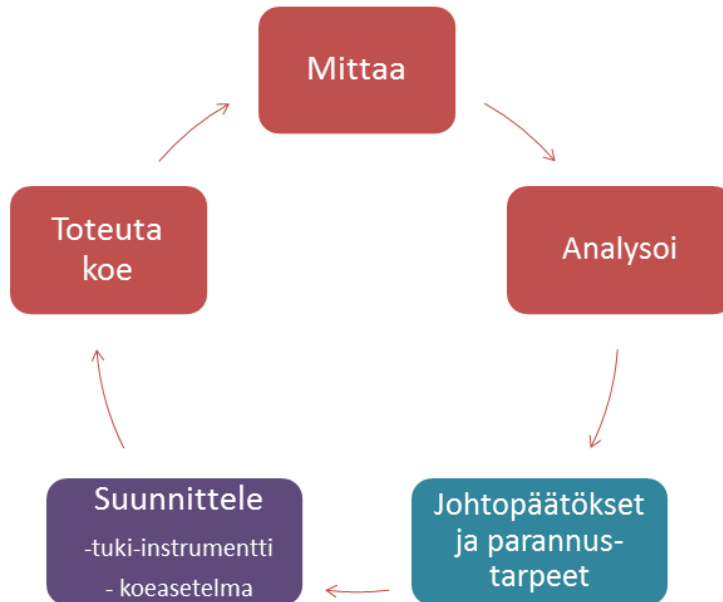
4.1 Vaikuttavuusarvioinnin kehitys-arviointi -sykli

Järjestelmällinen, säännöllisesti toteutettava yritystukiohjelmakohtainen satunnaistettu arviointitutkimus on luotettava tapa seurata yritystukijärjestelmän vaikuttavuutta. Se mahdollistaa myös tietopohjan ylläpidon. Ajantasaisen tiedon tärkeys yritystukijärjestelmän ohjaamisen kannalta on tärkeää, koska yrityssektori ja sen kansainvälinen toimintaympäristö ovat jatkuvan muutoksen alla. Tämä johtaa siihen, että aiemmin toimivaksi todetut tukimuodot eivät välttämättä ole enää nykyisin vaikutuksiltaan samanlaisia. Luotettava arviointi on myös hyödyllinen työkalu tukiohjelmista vastaaville, kun he haluavat seurata tukien tuloksellisuutta. Systemaattinen arviointi auttaa kehittämään yritystukien tehokasta kohdentumista ja mitoitusta.

Arvioinnin systemaattiseksi toteuttamiseksi yritystukiohjelmien tulisi seurata kehitys-arviointi-sykliä, kuviossa 3 hahmottelulla tavalla. Tässä tukiohjelma ensin suunnitellaan etukäteisarvion varassa mahdollisimman hyvin tavoitteitaan saavuttavaksi. Tämän jälkeen toteutetaan satunnaistettu arviointitutkimus vaikutusten selvittämiseksi. Tulosten pohjalta tukiohjelmaa voidaan suunnata uudelleen ja tuki-instrumenttia muuttaa vaikuttavampaan suuntaan. Aikajoin toteutettavalla syklillä voidaan varmistaa, että yritystukijärjestelmän tehokkuutta ei heikennä heikosti toimivat tuki-instrumentit. Luotavalla tietopohjalla voidaan odottaa olevan myös synergiaetuja, sillä tietoa hyvin vaikuttavien ohjelmien parhaista käytännöistä voidaan hyödyntää myös muissa ohjelmissa.

Arviointisyklin käyttö lyhyellä aikavälillä edellyttää suhteellisen nopeata vaikutusten syntymistä. Esimerkiksi pitkissä teknologian kehityshankkeissa lopulliset innovaatiotuotokset voivat syntyä pitkälläkin viiveellä. Tällaisissa tapauksissa on hyödyllistä arvioida ohjelman vaikutuksia lyhyellä aikavälillä epäsuorilla mittareilla. Tällaisia voivat olla esimerkiksi panosmuuttajat, joiden muutokset ovat tyypillisesti edellytyksenä pitkien vaikutusten syntymiselle. Esimerkki tästä olisi se, että innovaatiotukien kohdalla seurataan yritysten omia T&K-investointeja, vaikka pitkän aikavälin näkökulmasta kiinnostuksen kohteena olisivatkin innovaatiotuen avulla kehitetyt uudet teknologiat ja innovaatiot.

Kuvio 3: Kehitys-arviointi -sykli



4.2 Käytännön näkökohtia tuki-instituutioiden osallistamisesta

List (2011) ja Floyd ja List (2016) ovat pohtineet, mihin tekijöihin on perusteltua käytännössä kiinnittää huomiota, kun tutkijat toimivat yhteistyössä erilaisten organisaatioiden kanssa ja kun satunnaistettujen koejärjestelyiden avulla halutaan tutkia organisaatioihin liittyviä tai niitä kiinnostavia kysymyksiä. Heidän luetteloaan näistä tekijöistä voidaan soveltaa valikoidusti yritystukien arviointiin:

- *Talousteoria:* Yritystukiin liittyvät kysymykset ovat pitkälle taloudellisia ja niiden käyttäytymisvaikutukset syntyvät yritysten ja ihmisten taloudellisten valintojen seurauksena. Tästä syystä on perusteltua käyttää talousteoriaa ja taloustieteellistä tutkimusta hyväksi sekä koeasetelmaa suunniteltaessa että sen tuloksia analysoitaessa ja tulkittaessa (vrt. kuvio 3).
- *Institutionaaliset erityispiirteet:* Yritystukia ei myönnetä tyhjiössä vaan tiettyssä institutionaalisessa ympäristössä. Tästä syystä on tarpeen hyödyntää tarvittavaa markkina- ja tuki-instrumenttiosaamista, jotta toimintaympäristön erityispiirteet tulevat otetuksi tarkoituksenmukaisella tavalla huomioon, kun koejärjestelyä suunnitellaan, kun koetta toteutetaan, kun tuloksia mitataan ja kun niitä analysoidaan.

- *Otoskoko ja vertailuryhmä:* Satunnaistamisen tärkein tehtävä on muodostaa verrokkiryhmä, johon toimenpidettä ei kohdisteta, mutta joka on keskimääräisiltä ominaisuuksiltaan samankaltainen kuin toimenpiteelle altistettava koeryhmä. Tästä syystä on tärkeää varmistaa, että tutkimusasetelma tuottaa tarkoituksenmukaisen vertailuryhmän. Myös otoskoon on oltava riittävän suuri tilastollisen päättelyn näkökulmasta (ts. jotta voidaan sulkea sattuman vaikutus pois riittävän luotettavasti).
- *Yhteistyö ja osallistaminen:* Kokeiden käytännön järjestelyihin liittyy lukuisia yksityiskohtia, jotka riippuvat ohjelmien tavoitteista, institutionaalisista erityispiirteistä ja tarvittavista aineistoista. Kokeiden suunnittelu edellyttää siksi aina tiivistä yhteistyötä tutkijoiden ja ohjelmantoteuttavien tahojen välillä. On siksi tärkeää, että yritystukea myöntävän organisaation eri tasot tukevat kokeen toteuttamista ja ymmärtävät sen tavoitteet ja että organisaatio voi tuoda osaamistansa hankkeeseen.
- *Kustannukset ja hyödyt:* Koejärjestelyn kustannukset ovat usein näkyviä, mutta hyödyt eivät. On tarpeen varmistaa, että odotetut hyödyt ymmärretään, samoin kuin se, että osa hyödyistä voi realisoitua vasta viipeellä, esimerkiksi silloin kun tukea tai instrumenttia uudistetaan.
- *Tutkijoiden tavoitteet ja työnjako:* Tutkijat tekevät tutkimuksia yritystuista, koska niiden tuloksilla voidaan kehittää yritystukijärjestelmää ja tuottaa tietoa sen vaikuttavuudesta. Tutkijoiden yhtenä päämääränä on kuitenkin myös tieteellisen julkaisun tuottaminen, koska sen avulla voidaan varmistaa, että tulokset ovat luotettavia ja läpäisevät kriittisen tieteellisen arvioinnin. Olisikin hyvä sopia selvästi ja etukäteen, että tutkimuksen tulokset voidaan julkaista riippumatta siitä, mitä ne ovat. Tutkijoiden erityisvahvuus on erilaisten tutkimusasetelmien ymmärrys ja menetelmäosaaminen. Hyväkään tutkijaryhmä ei kuitenkaan tiedä kaikkia vastauksia etukäteen ja kun koejärjestelyä suunnitellaan ja toteutetaan, ongelmia syntyy. Ne on tehty yhteistyössä ratkaistaviksi.

List (2011) ja Floyd ja List (2016) korostavat myös, että on perusteltua hyvissä ajoin varmistaa, miten tutkimukseen mahdollisesti liittyvät eettiset kysymykset ratkaistaan ja mikä organisaatiotaho myöntää luvan.

4.3 Muut ekonometriset arviointimenetelmät

Entä mitä voidaan tehdä silloin, kun satunnaistaminen ei ole mahdollista? Tässä kohtaa on perusteltua todeta se, että on sekä tiettyjä yhteiskunnallisista periaatteista kumpuavia esteitä että käytännöllisiä rajoitteita, joiden vuoksi satunnaistamista ei voida käytännössä joidenkin tukityyppien kohdalla hyödyntää. Tämä ei kuitenkaan tarkoita, että edellä kuvattua arviointisykliä ei voisi käytännössä soveltaa.

Mikäli todetaan, että jonkin yritystuen vaikuttavuusarvioinnissa ei voida hyödyntää satunnaistamista (ks. myös tämän raportin johtopäätösluku), vaikka ko. tuki periaatteessa olisi perusteltua saattaa järjestelmällisen vaikuttavuusarvioinnin piiriin, on perusteltua harkita, voidaanko vaihtoehtoisesti hyödyntää muita tehokkaiksi tiedettyjä ekonometrisia vaikuttavuusarviointimenetelmiä. Myös näiden arvioiden luotettavuutta voidaan parantaa, jos yritystukiohjelmat seuraavat kehitys-arviointi-sykliä ja jos ekonometrinen vaikuttavuusarviointi otetaan jo suunnitteluvaiheessa huomioon ja osaksi ko. sykliä.

Yksi vaikuttavuusarviointiin soveltuvista ekonometrisista menetelmistä on ns. regressioepäjatkuvuusmenetelmä (*regression discontinuity design*, RD-asetelma; ks. mm. Hahn, Todd, ja van der Klaauw 2001 ja Lee 2008):

- RD-asetelmaa voitaisiin hyödyntää varsin suoraviivaisesti yritystukien vaikuttavuuden arvioinnissa, mikäli yritysten tukikelpoisuus pisteytetään ja mikäli tuki myönnetään (todennäköisemmin) tietyn pistemäärän ylittäville, mutta ei sen alle jääville yrityksille. RD-asetelmalla voidaan arvioida tuen vaikuttavuutta vertaamalla tukea saaneita yrityksiä niihin yrityksiin, jotka eivät saaneet tukea ja jotka juuri ja juuri jäivät tukeen oikeuttavan pistemäärän alle (Hahn, Todd, ja van der Klaauw 2001). Mikäli yritykset eivät voi vaikuttaa (manipuloida) omaan pistemääräänsä, voidaan osoittaa, että tuki kohdentuu kynnysen lähellä ikään kuin satunnaisesti (Lee 2008). Tämä tekee tuen vaikuttavuuden luotettavan arvioinnin mahdolliseksi.
- RD-asetelmaa on jo hyödynnetty varsin monissa yritystukia arvioineissa kansainvälisissä tutkimuksissa sekä suorien tukien (ks. mm. Bronzini and Iachini 2014; Howell 2017) että veromuotoisten tukien (ks. mm. Dechezleppre ym. 2016) osalta. Esimerkiksi Howellin tutkimus osoittaa, että ne nuoret yritykset, jotka saivat yhdysvaltalaisen T&K-tukiohjelman (*SBIR grant program*) puitteissa julkista rahoitusta, kasvattivat merkittävästi todennäköisyyttä saada myöhemmin yksityistä pääomarahoitusta. Tämä tulos perustuu siihen, että ko. ohjelmasta tukea saaneita verrattiin tukea hakeneisiin mutta pistemäärässä juuri rahoitusrajan alapuolelle jääneisiin yrityksiin. Tällä menetelmällä havaittiin myös, että ko. rahoitusta saaneet yritykset kasvoivat myös nopeammin ja patentoivat enemmän.
- Uusimpien RD-menetelmien on havaittu kykenevän toistamaan tuloksia, jotka on saatavissa satunnaistetuilla koejärjestelyillä. Tällaisia tuloksia on saatu myös suomalaisella aineistolla (ks. Hyytinen, Meriläinen, Saarimaa, Toivanen ja Tukiainen 2018).

RD-menetelmien tehokas hyödyntäminen edellyttää käytännössä sitä, että tukiohjelmaa suunniteltaessa otetaan huomioon myöhempi vaikuttavuusarvion tekotarve. Perusidea on se, että sen sijaan, että tuen myöntövaiheessa sitä myönnetään satunna-

sesti, tukea kohdennetaan pisteytykseen perustuen: Mikäli tukea myönnetään todennäköisemmin tietyn pistemäärän ylittäville, mutta ei sen alle jääville yrityksille, voidaan RD-menetelmää tyypillisesti hyödyntää.

Toisinaan tukien vaikuttavuutta voi olla perusteltua myös ns. rakenteellisilla ekonometrisilla malleilla (*structural econometric models*; ks. esimerkiksi Takalo, Tanayama, ja Toivanen 2013). Näiden mallien etu on mm. se, että niiden avulla voidaan selvittää myös sellaisien (uusien) tukimuotojen ja -instrumenttien toimintaa, joita ei vielä ole käytännössä kokeiltu. Ne myös voivat tuottaa uutta ja tarkempaa tietoa siitä, miksi tuet toimivat ja ketkä niistä erityisesti hyötyvät.

Lisäksi kyseeseen saattaa tulla ns. synteettisen kontrollimenetelmän (*synthetic control method*) hyödyntäminen (ks. esimerkiksi Abadie, Diamond ja Heilmüller 2010). Tämä menetelmä sopii mahdollisesti tilanteisiin, joissa tukea saavien yritysten lukumäärä on pieni ja joissa satunnaistaminen ei siksi ole mahdollista. Palaamme tähän menetelmään lyhyesti johtopäätösluvussa.

5 Johtopäätökset ja loppupohdintaa

5.1 Minkälaisia yritystukia olisi perusteltua arvioida satunnaistettujen koejärjestelyiden avulla?

Tilastokeskuksen yritystukitilaston mukaan Suomessa sai vuonna 2017 yritystukea (sis. suorat tuet, lainat ja takaukset) 28412 yritystä. Näistä alkutuotannossa toimi 770, teollisuudessa 3770, rakentamisessa 2510, kaupan alalla 3886, palvelualoilla 9833 ja muilla toimialoilla 1417 yritystä. Lisäksi tukea sai 6226 yritystä, joiden toimiala on tuntematon ja jokin määrä yrityksiä, joilla ei ole y-tunnusta alkuperäisaineistossa (Tilastokeskus 2018). TEM:in (2017) julkaisun mukaan euromääräisesti arvioituna suurimmat tuet ovat energiaintensiivisten yritysten veronpalautus (pl. kaivostoiminta), uusiutuvan energian tuotantotuki (tuulivoima, biokaasu, puupolttoaine, metsähake), maatalousyrittäjien, turkistuottajien ja poronhoitajien lomituspalvelujen rahoitus, TKI –toiminnan tukeminen (Tekesin / Business Finlandin avustukset yrityksille) ja Lainat TKI-toimintaan (Tekes / Business Finland).¹⁰

Yritystukien vaikuttavuusarvioinnin keskeisiä haasteita on, että tietyissä tuki-instrumenteissa tukea saavien yritys joukko on määrällisesti melko pieni. Lisäksi joissakin tuissa tuen saajien kokojakauma on vino. Edellinen viittaa siihen, että tiettyjä tukia myönnetään vain melko harvoille yritykselle ja jälkimmäinen siihen, että merkittävä osa tukisummasta kohdentuu muutamalle isolle yritykselle ja iso joukko pienempiä yrityksiä saa pienen osuuden kokonaistukisummasta.

Yleisellä tasolla – ja myös päätöksentekoteoreettisesti perustellusti – voidaan sanoa, että satunnaistetut vertailukokeet ovat erityisen hyödyllisiä (ks. Banerjee ja muut 2017),

- kun yritystuen vaikuttavuusarviolla pyritään näyttämään jollekin kohderyhmälle tai vakuuttamaan laajempi yleisö siitä, miten hyvin ko. tuki-instrumentti toimii (tai ei toimi); ja
- kun asetelma on sellainen, että on mahdollista tutkia riittävän suurta koe- ja vertailuaineistoa.

¹⁰ On rajanvetokysymys, onko STM:n alaisuuteen kuuluva lomituspalvelu yritystukea vai osa sosiaaliturvaa. Julkisesti rahoitetun lomituspalvelun tarkoitus on tukea yrittäjän sosiaaliturvan toteutumista, työssä jaksamista ja työurien pidentämistä. Valtiontuki myönnetään tämän tuen tapauksessa henkilölle, joka toimii yrittäjänä tietyllä sektorilla.

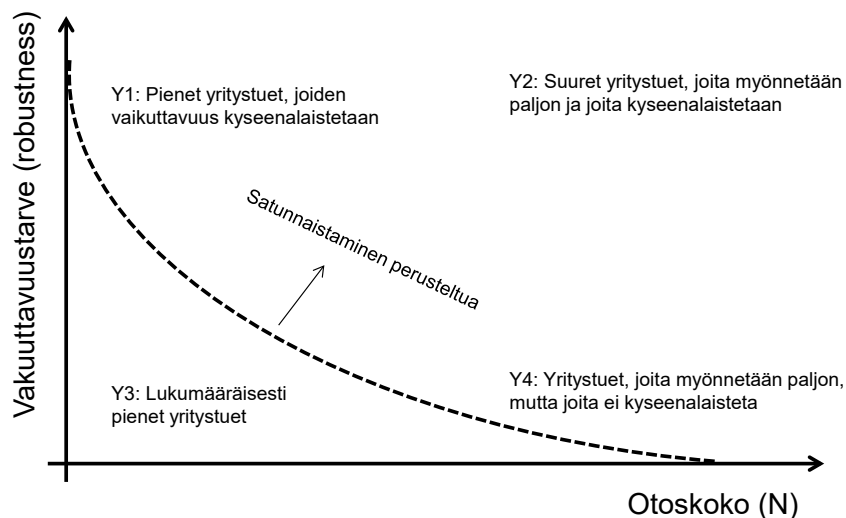
Vakuuttavuustarve liittyy tässä siihen, että eri sidosryhmillä ja toimijoilla, kuten poliittisilla päättäjillä ja äänestäjillä, voi olla hyvin monenlaisia etukäteisnäkemyksiä (*prior beliefs*) siitä, miten erilaiset yritystuet vaikuttavat. Mitä erilaisempia etukäteisnäkemykset ovat, sitä perustellumpaa on toteuttaa satunnaistettuihin vertailukokeisiin perustuva selvitys siitä, miten erilaiset yritystuet toimivat ja mikä toimenpide olisi vaikuttavin. Palaamme tähän uudelleen alla.

Alla oleva Kuvio 4 perustuu Banerjee ja muut (2017) tutkimukseen ja se havainnollistaa edellä mainittua valintatilannetta: Kuviossa pystyakselilla mitataan sitä, kuinka suuri ulkopuolisten vakuuttavuustarve on. Vaaka-akseli puolestaan kuvaa mahdollista otoskokoja, joka olisi käytettävissä, kun toimenpiteen vaikuttavuutta arvioidaan.

Kuvioon on lisätty neljä erilaista yritystukimuotoa, Y1, Y2, Y3, Y4. Näistä ensimmäinen, Y1, viittaa yritystukiin, joissa saajia on vähän mutta joiden toimivuudesta saattaisi olla tarvetta tuottaa näyttöä laajemmalle yleisölle ja eri sidosryhmille. Yritystuki Y2 viittaa yritystukiin, joissa saajia on lukumääräisesti paljon ja joiden toimivuudesta saattaisi olla tarvetta tuottaa näyttöä laajemmalle yleisölle ja eri sidosryhmille. Yritystuet Y3 ja Y4 ovat puolestaan sellaisia, että niiden toimivuudesta ei koeta olevan tarvetta tuottaa näyttöä laajemmalle yleisölle ja eri sidosryhmille. Tukien Y3 ja Y4 ero on se, että Y3 on lukumääräisesti pieni tuki, mutta Y4 kohdentuu suurelle joukolle yrityksiä.

Tähän kuvioon palataan vielä jäljempänä, mutta sen perusidea on osoittaa, minkälaiset tuet tai toimet on ainakin perusteltua altistaa satunnaistettuihin koejärjestelyihin perustuvilla vaikuttavuusarvioinneilla. Kuvio kertoo, että tukien Y1, Y2 ja Y4 kaltaisten tuki-instrumenttien arvioinnissa on todennäköisesti perusteltua hyödyntää satunnaistamista.

Kuvio 4: Satunnaistettuihin koejärjestelyihin perustuvien vaikuttavuusarviointien tarve



Lähde: Kirjoittajat (mukaien Banerjee ja muut 2017, kuvio 1).

5.2 Kohti järjestelmällistä vaikuttavuusarviointia?

Mielestämme yritystukien vaikuttavuuden arviointi olisi syytä systymatisoida Suomessa nykyistä paremmin. Onkin hienoa, että Valtioneuvosto asetti 28.3.2019 yritystukien tutkimusjaoston. Kyseessä on riippumaton tutkimusjaosto, joka toimii Työ- ja elinkeinoministeriön yritystukineuvottelukunnan yhteydessä ja jonka tehtävänä on arvioida yritystukijärjestelmää. Tutkimusjaosto laatii vuosittain arvio yritystukijärjestelmän kehittymisestä, kehittää vaikuttavuuden tutkimusmetodologiaa ja tekee riippumattomia arvioita yritystuista.

Alla tehtävät ehdotukset ovat kirjoittajien näkemyksiä siitä, miten satunnaistettujen koejärjestelyiden hyödyntämistä vaikuttavuusarvioinnissa voitaisiin jatkossa konkreettisesti edistää:

Ehdotus 1: On tarpeen pyrkiä identifioimaan, mitkä Suomessa käytössä olevista tuki-instrumenteista ovat Kuviossa 4 mainittujen tukien Y1, Y2 ja Y4 kaltaisia. Kuvioon 4 perustuva luokittelu tarjoaisi pohjan mahdolliselle myöhemmälle päätökselle altistaa tietyt tuet nykyistä järjestelmällisemmän, satunnaistamista hyödyntävän vaikuttavuusarvioinnin piiriin.

- *Vakuuttavuustarpeen arviointi:* Kuten edellä todettiin, satunnaistetut vertailukokeet ovat perusteltuja, kun vaikuttavuusarviolla on identifioitu tarve pyrittään vakuuttamaan laajempi yleisö tai päätöksentekijät siitä, miten hyvin tuki-instrumentti toimii (tai ei toimi) ja kun asiasta vallitsee erilaisia näkemyksiä. Tämän kirjoituksen puitteissa ei ollut mahdollista arvioida sitä, mitkä ovat tällaisia tuki-instrumentteja. Käytännössä yhtenä hyvänä lähtökohtana vakuuttavuustarpeen arvioinnille voisi toimia kuitenkin ns. yritystukien liikennevalot, jotka on julkaistu TEM (2017) julkaisun liitteessä 3. Lisäksi, mikäli näkemyserot ovat erityisen suuria silloin, kun kyse on euromääräisesti suurista yritystuista, vakuuttavuustarve voi olla suoraan verrannollinen tuen kokonaissummaan.
 - Vakuuttavuustarve voi myös muuttua ajan myötä. Jonkin tuen vakuuttavuustarve voi olla esimerkiksi alkujaan pieni, jos kyseessä on tuki, josta on hiljattain voitu tehdä hyvä näennäiskoeasetelmaan perustuva vaikuttavuusarvio ja jos myös muu (kansainvälinen) tutkimusnäyttö tukee arviosta saatuja johtopäätöksiä. Koska tuen kohdentuminen on ohjelmakohdasta ja koska tukea hakevien ja saavien yritysten toimintaympäristö muuttuu, vakuuttavuustarve riippuu myös siitä, miten paljon vaihtelua toimintaympäristössä on.

- *Yritystuen suuruuden arviointi (lukumäärä, N):* Satunnaistettu koejärjestely edellyttää, että on mahdollista tutkia riittävän suurta koe- ja vertailuaineistoa. Tämä on kuitenkin konteksti- ja tukisidonnainen, koska riittävän suuri N riippuu mm. siitä, mihin vaikuttavuutta koskeviin kysymyksiin halutaan vastaus, siitä, minkälainen arviointitilanne kohdataan (vrt. myös alaluvut 3.1-3.3) ja siitä, minkä verran tulemamuuttujissa on vaihtelua (vrt. alaluku 2.2. kokeen koosta ja voimasta).
 - Yksiselitteistä vastausta siihen, mikä on riittävän iso N, jotta satunnaistamista voidaan tehdä, ei siis ole. Viitaten kuvioon 4, on tapauskohtaisesti ja soveltuvia tilastoaineistoja hyödyntäen arvioitava, voidaanko ja kannattaako yritystukityyppiä Y1 oleva tuki saattaa satunnaistetun vaikuttavuusarvioinnin piiriin.
 - Käytössä olevat julkaistut tilastotiedot eivät ole riittävän yksityiskohtaisia, jotta niiden avulla voitaisiin riittävän tarkasti yksilöidä, kuinka paljon tuen saajia on tuki-instrumentia kohden, tukiohjelmittain ja/tai esimerkiksi alueittain tai toimialoittain. Tämänkaltaisen tieto on kuitenkin monien tukien osalta saatavissa Tilastokeskuksen yritystukitietokannasta.

Tarkoitamme yritystukien systemaattisella arvioinnilla ensinnäkin sitä, että kullekin tuelle tai tukimuodolle olisi laadittava etukäteen arviointisuunnitelma, joka kuvaa, miten tuen vaikuttavuutta tullaan jatkossa arvioimaan (vrt. myös alaluvun 4.1 keskustelu). Osana tätä suunnitelmaa määriteltäisiin, missä määrin jo tuen myöntämisympäristössä on tarpeen tehdä toimia – kuten esimerkiksi tuen myöntömenettelyyn liittyvää satunnaistamista – jotka tekevät luotettavan ex post vaikuttavuusarvioinnin mahdolliseksi. Arviointisuunnitelmassa olisi perusteltua kuvata myös mm. se, miten vaikuttavuusarviointia varten kerätään tilastoaineistoa tukea hakeneista ja saaneista yrityksistä. Oleellinen osa tällaista järjestelmällistä vaikuttavuusarviointia on se, että arviointitulosten ansiosta tukien vaikuttavuutta voidaan tehostaa.

Kuten tässä raportissa on jo aiemmin todettu, satunnaistaminen ei sovi käytännön syistä kaikkien yritystukien arviointimenetelmäksi. Esimerkiksi lukumääräisesti pienet tuet, joiden tehokkuutta ei kyseenalaisteta (kuviossa 4 tuen Y3 kaltaiset tuet), saattaisi olla perusteltua jättää laajamittaisemman, satunnaistettuja koejärjestelyitä hyödyntävän vaikuttavuusarvioinnin ulkopuolelle. Tähän liittyen on kuitenkin tarpeen esittää seuraava varaus: Koska tukea saavien yritysten lukumäärällä pieneksi arvioitu tuki voi olla yhteiskunnallisesti merkittävä, on erikseen pohdittava, miten sellaisia tukimuotoja on tarkoituksenmukaista arvioida. Tukien Y3 osalta kyseeseen saattaa tulla (aiemmin jo lyhyesti mainitun) synteettisen kontrollimenetelmän hyödyntäminen (Abadie, Diamond ja Heilmüller 2010). Tätä menetelmää voidaan tietysti varauksin soveltaa tilanteissa, joissa satunnaistaminen ei ole pienen N:n vuoksi mahdollista. Se siis soveltuu tukiin, joissa on vain pienehkö määrä kooltaan isoja tuen saajia. Näille isoille yrityksille on tyypillisesti vaikea löytää verrokkeja, mutta synteettisen kontrollimenetelmän idea on, että verrokki voidaan tuottaa synteettisesti tilastoaineistoon perustuen. Yritystukien

tapauksessa tämä voisi tarkoittaa mm. kotimaisen ja kansainvälisen yritysaineiston yhdistelyä joltakin tietyltä toimialalta.

Ehdotus 2: Järjestelmällistä vaikuttavuusarviointia on syytä kehittää vaiheittain. Ehdotamme tästä syystä kahden pilottiprojektin käynnistämistä, joissa suunnitellaan yhteistyössä relevanttien ministeriöiden, Yritystukineuvottelukunnan ja Yritystukien tutkimusjaoston ja tukia myöntävien organisaatioiden kanssa, kuinka satunnaistamista sisältäviä koejärjestelyitä voitaisiin hyödyntää tiettyjen tukien arvioinnissa:

- *Pilotti 1:* Kohteeksi valitaan yhteiskunnallisesti tärkeäksi koettu – kuten esimerkiksi tuen kokonaissumman vuoksi merkittäväksi arvioitu – ja tuen saajien suuren lukumäärän vuoksi suureksi luettava yritystukimuoto. Tällainen tuki voisi esimerkiksi olla Business Finlandin (Tekesin) myöntämä innovaatorahoitus: Vuonna 2017 Tekesin innovaatorahoitus yrityksille ja tutkimukseen oli yhteensä 466,2 milj. euroa. Yrityksille suunnatusta innovaatorahoituksesta pk-yritysten osuus oli 76%. Vuonna 2017 rahoitusta myönnettiin 1 679 yritykselle ja maksettiin 3 019 yritykselle, johtuen mm. siitä että tukea voidaan maksaa useana vuotena myönnön jälkeen. Kuviossa 4 nämä tuet voivat olla joko tukien Y2 tai Y4 tyyppisiä. Innovaatorahoituksen keskeinen tavoite on vahvistaa yritysten innovointikykyä ja siten pitkän aikavälin talouskasvun edellytyksiä. Vaikka yleisesti innovaatorahoitukselle asetettuja tavoitteita ja niiden toimivuutta ei kyseenalaisteta, voidaan silti perustellusti esittää kysymys, kuinka tehokkaita tuet ovat (eli mikä on niiden vaikuttavuuden suuruusluokka) ja kuinka hyvin erilaiset suunnatut tai valikoivat ohjelmat toimivat. Ylipäätään on vähän tutkimustietoa siitä, miten eri innovaatorahoituksen muodot ja instrumentit vertautuvat vaikuttavuudeltaan toisiinsa ja kuinka niiden toimivuutta voitaisiin nykyisestä vielä tehostaa. On luontevaa, että ensisijaisesti pilottihankkeessa selvitetään, kuinka innovaatorahoitus vaikuttaa yritysten omarahoitteiseen T&K-toimintaan ja muihin mitattavissa oleviin innovointipanostuksiin. Tämän lisäksi pilotissa olisi perusteltua selvittää, kuinka innovaatorahoitus vaikuttaa innovaatiotoiminnan tuotoksiin, kuten esimerkiksi uusien innovaatioiden määrään, niiden luonteeseen (prosessi- vs. tuoteinnovaatio; teknologisen kehityksen suunta), yritysten kokonaistuottavuuteen ja asemaan kansainvälisillä markkinoilla. Lisäksi teknologiatyyppin mukaan suunnatuissa ohjelmissa voidaan myös tarkastella vaikutuksia patentointiin tietyissä relevanteissa teknologia-alueissa (ks. esim. Koski ym, 2019).

- Pilotti 2:* Kohteeksi valitaan yritystukien mahdollisia pitkän aikavälin tuottavuusvaikutuksia ja vaikutusmekanismeja paljastava tukimuoto. Tällä tarkoitetaan joko sellaisia tukimuotoja, jotka saattavat ei-toivottujen vaikutustensa vuoksi nakertaa yrityssektorin pitkän aikavälin tuottavuutta ja/tai sellaisia yritystukimuotoja, joilla saattaa olla negatiivisia markkinavaikutuksia. Tällainen tuki voisi esimerkiksi olla yritysten palkkatuki, jota myöntävät TE-toimistot. Palkkatuen ensisijainen tavoite on yhteiskunnallinen: sillä pyritään integroimaan työtön työnhakija takaisin työmarkkinoille, kun hänellä on puutteita ammatillisessa osaamisessa tai tehtävissä selviytymiseen vaikuttava vamma tai sairaus. Palkkatuki voi olla enintään 30%, 40% tai 50% palkkauskustannuksista, enimmäismäärän ollessa 1400 euroa kuukaudessa. Tuki-instrumenttia varten on varattu vuodelle 2019 57 milj. euron määräraha. Enimmäismääräkorvauksella laskettuna tämä vastaa noin 40700:ää tuettua työkuukautta. Palkkatuki sijoittuu kuviossa tukityyppien Y1 ja Y4 väliin: kyseessä on keski-suuri yritystukimuoto, joilla voi olla kielteisiä markkinavaikutuksia työmarkkinoiden kautta ja jotka voivat olla tuottavuuskasvua heikentäviä. Palkkatuen kohdalla olisi ensisijaisen tärkeää selvittää sitä, syrjäyttääkö tuetun työntekijän palkkaaminen yrityksessä ilman palkkatukea työskentelevää työvoimaa. Tämän lisäksi olisi perusteltua selvittää pidemmän aikavälin vaikutuksia, eli sitä, vaikuttaako tukityöjakso tuetun työntekijän työmarkkinasemaan jakson jälkeen. Lisäksi olisi hyödyllistä selvittää, miten palkkatuki vaikuttaa palkkaavan yrityksen tuottavuuskehitykseen, liikevaihtoon, kannattavuuteen ja markkina-asemaan. Samoin on tarpeen selvittää, miten palkkatuki vaikuttaa tehottomien yritysten selviämistodennäköisyyteen ja laajenemiseen, ja tätä kautta koko talouden tuottavuutta kohottavaan yritys- ja työpaikkarakenteiden muutokseen.

Ehdotuksemme on, että *Yritystukineuvottelukunnan yhteydessä toimiva Yritystukien tutkimusjaosto* koordinoi yllä kuvattujen pilottihankkeiden käynnistämisen ja kutsuu kokoon ryhmän, joka konkreettisesti suunnittelee pilottihankkeiden puitteissa tehtävät vaikuttavuusarviot. Mukaan ryhmään on syytä kutsua relevantin ministeriön edustaja, alan tutkijoita ja tukia myöntävien organisaatioiden edustajia. Tavoitteena voisi olla, että tämä työ käynnistyy syksyllä 2019 tai viimeistään vuoden 2020 alussa. Näiden pilottihankkeiden puitteissa määritellään tarkemmin, miten ne toteutettaisiin. Tarkoituksena ei esimerkiksi ole, että koko Business Finlandin myöntösumma (466 miljoona €) olisi pilottisatunnaiskokeessa mukana. Se, mitä tuki-instrumenttia satunnaiskoe voisi koskea, selvitetään osana pilottihanketta.

Piloteista saatujen kokemusten pohjalta voidaan järjestelmällistä vaikuttavuusarviointia sitten laajentaa muihin tukimuotoihin. Tavoitteeksi on hyvä asettaa se, että kullekin tuelle tai tukimuodolle on vuoteen 2025 mennessä laadittu arviointisuunnitelma, joka kuvaa, miten tuen vaikuttavuutta tullaan arvioimaan.

Lähteet

- Abadie, A.; Diamond, A.; ja Hainmüller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's Tobacco Control Program *Journal of the American Statistical Association*, 105 (490), 493–505.
- Angrist, J. D. (2014). The perils of peer effects. *Labour Economics*, 30, 98-108.
- Athey, S. ja Imbens, G., 2017, The econometrics of randomized experiments, teoksessa: *Handbook of Field Experiments Volume 1* (toimittaneet: A V. Banerjee ja E. Duflo), North-Holland, luku 3, 75-140.
- Bakhshi, H., Edwards, J. S., Roper, S., Scully, J., Shaw, D., Morley, L., & Rathbone, N., 2015, Assessing an experimental approach to industrial policy evaluation: Applying RCT+ to the case of Creative Credits. *Research Policy*, 44(8), 1462-1472.
- Banerjee, A., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukerji, S., Shotland, M., ja Walton, M., 2017, From proof of concept to scalable policies: Challenges and solutions, with an application." *Journal of Economic Perspectives*, vsk. 31(4), 73-102.
- Banerjee, A., Chassang, S. ja Snowberg, E., 2017, Decision theoretic approaches to experiment design and external validity, teoksessa: *Handbook of Field Experiments Volume 1* (toimittaneet: A V. Banerjee ja E. Duflo), North-Holland, luku 4, 142-174.
- Berry, J., Fischer, G. ja Guiteras, 2018, Eliciting and utilizing willingness-to-pay: Evidence from field trials in Northern Ghana, *Journal of Political Economy*, tulossa.
- BIS, 2014, *Growth Vouchers programme: trial protocol (full)*. UK Department for Business, Innovation & Skills.
- BIS, 2016, *Growth Vouchers programme evaluation cohort 1: impact at six months*. BIS Research Paper No. 259. UK Department for Business, Innovation & Skills.
- Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., and Roberts, J., 2013, Does management matter? Evidence from India." Q 128 (February): 1–51.
- Bronzini, R. ja Iachini, E., 2014, Are incentives for R&D effective? Evidence from a regression discontinuity approach, *American Economic Journal: Economic Policy*, vsk. 6(4), 100–134.
- Bruhn, M., Karlan, D., & Schoar, A., 2018, The impact of consulting services on small and medium enterprises: Evidence from a randomized trial in Mexico. *Journal of Political Economy*, 126(2), 635-687.
- Chassang, S., Miquel, G. P. ja Snowberg, E., 2012, Selective trials: A principal-agent approach to randomized controlled experiments, *American Economic Review*, vsk. 102(4), 1279-1309.
- Deaton, A., 2010, Instruments, randomization, and learning about development, *Journal of Economic Literature*, vsk. 48(2), 424-55.
- Dechezleprêtre, A., Einiö, E., Martin, R., Nguyen, K.-T. ja Van Reenen, J. 2016, Do Tax Incentives for Research Increase Firm Innovation? An RD Design for R&D. NBER Working Paper No. 22405, July 2016.
- Einiö, E. 2014, R&D subsidies and company performance: Evidence from geographic variation in government funding based on the ERDF population-density rule. *Review of Economics and Statistics*, 96(4), 710-728.
- Fairlie, R. W., Karlan, D., & Zinman, J., 2015, Behind the GATE experiment: Evidence on effects of and rationales for subsidized entrepreneurship training. *American Economic Journal: Economic Policy*, 7(2), 125-61.
- Floyd, E. ja List, J. A., 2016, Using field experiments in accounting and finance, *Journal of Accounting Research*, vsk. 54(2), 437-475.
- Glennerster, R. & Takavarasha, K. 2013. "Running randomized evaluations : a practical guide." Princeton, New Jersey : Princeton University Press.
- Hahn, J., Todd, O. ja van der Klaauw, W., 2001, Identification and estimation of treatment effects with regression-discontinuity design, *Econometrica*, vsk 69(1), 201–209.
- Hall, B., Griliches, Z. and Hausman, J. 1986, Patents and R&D: Is There A Lag? *International Economic Review*, 27(2) 265-284.
- Halloran, M. E. ja Hudgens, M. G., 2016, Dependent happenings: A recent methodological review, *Current Epidemiology Reports*, vsk. 3(4), 297-305.

- Hammar, M. Hyytinen, A. ja Tuimala J., 2018. Enhancing the use of digital government services: Evidence from a natural field experiment, käsikirjoitus.
- Hanna, R., Duflo, E., ja Greenstone, M., 2016, Up in smoke: The influence of household behavior on the long-run impact of improved cooking stoves, *American Economic Journal: Economic Policy*, vsk. 8(1), 80-114.
- Heckman, J. J., & Smith, J. A. 1995, Assessing the case for social experiments. *Journal of Economic Perspectives*, 9(2), 85-110.
- Howell, S. T., 2017, Financing innovation: Evidence from R&D grants, *American Economic Review* 107 (4), 1136-1164.
- Hyytinen, A. ja Toivanen, O., 2005, Do financial constraints hold back innovation and growth? Evidence on the role of public policy, *Research Policy*, 34(9), 1385-1403.
- Hyytinen, A., Meriläinen, J., Saarimaa, T., Toivanen, O. ja Tukiainen, J., 2018, When does regression discontinuity design work? Evidence from random election outcomes, *Quantitative Economics*, 9(2), 1019-1051
- Hämäläinen, Kari, Kanninen, Ohto; Simanainen, Miska; Verho, Jouko (2019). Perustulokokeilun ensimmäinen vuosi. VATT Muistiot 56. Valtion taloudellinen tutkimuskeskus.
- Hämäläinen, K., Uusitalo, R., Vuori, J., 2008, Varying biases in matching estimates: Evidence from two randomised job search training experiments, *Labour Economics*, 15(4), 604-618.
- Imai, K., Tingley, D., & Yamamoto, T. (2013). Experimental designs for identifying causal mechanisms. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 176(1), 5-51.
- Koski, H., Ollikka, K. ja Ylhäinen, I., 2019, Environmental policy, green innovation and market developments. Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja 2019:36.
- Kosonen, T. ja Ropponen, O., 2015, The role of information in tax compliance: Evidence from a natural field experiment, *Economics Letters*, 129(C), 18-21.
- Laukkanen, M. ja Maliranta, M., 2019, Yritystuet ja kilpailukyky. Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja 2019:33.
- Laukkanen, M., Ollikka, K. ja Tamminen, S., 2019, The impact of energy tax refunds on manufacturing firm performance: evidence from Finland's 2011 energy tax reform, Valtioneuvosto, Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja No. 2019:32.
- Lee, D. S., 2008, Randomized experiments from a non-random selection in U.S. House Elections, *Journal of Econometrics*, vsk. 142(2), 675-697.
- List, J. A., 2011, Why economists should conduct field experiments and 14 tips for pulling one off, *Journal of Economic Perspectives*, vsk. 25, 3-15.
- Manski, C. F., 1993, Identification of endogenous social effects: The reflection problem, *Review of Economic Studies*, vsk. 60(3), 531-542.
- Sacerdote, B., 2014, Experimental and quasi-experimental analysis of peer effects: Two steps forward? *Annual Review of Economics*, vks. 6(1), 253-272.
- Olken, B. A., 2015, Promises and perils of pre-analysis plans. *Journal of Economic Perspectives*, 29(3), 61-80.
- Takalo, T., Tanayama, T. ja Toivanen, O., 2013, Estimating the benefits of targeted R&D subsidies, *Review of Economics and Statistics*, 95(1), 255-272.
- TEM, 2017, *Virkamiesselvitys yritystuista ja niiden vaikutuksista*, Työ- ja elinkeinoministeriön julkaisuja, Yritykset, 22/2017.
- Tilastokeskus, 2018, *Suomen virallinen tilasto (SVT): Yritystukitilasto 2017*, Helsinki: Tilastokeskus
- Vasquez-Bare, G., 2018, Identification and estimation of spillover effects in randomized experiments, manuscript, University of California, Santa Barbara.
- Vivalt, E., 2015, Heterogenous treatment effects in impact evaluation, *American Economic Review, Papers & Proceedings*, vsk. 105(5), 467-470.

TIETOKAYTTOON.FI

