

LIITE A.1 – Samankaltaistaminen

Kaltaistamisen oletuksena on oletus yhteisestä tuesta (engl. *common support*), jolla tarkoitetaan sitä, että yksikön ominaisuuksien perusteella ei tule voida suoraan ennustaa, päätyykö yksikkö käsittelyn kohteeksi (Caliendo & Kopeinig 2005, 4). Kun X kuvaa havaittavissa olevia taustamuuttujia ja T on kokeiluun osallistumisen indikaattori, oletus voidaan ilmaista matemaattisesti niin, että

$$0 < P(T = 1|X) < 1.$$

Kokeiluun osallistuneissa ja osallistumattomissa kunnissa tulisi siis olla muuttujien X arvoissa päällekkäisyyttä. Mikäli edellä mainittu oletus ei täyty, menetelmä hylkää tällaiset yksiköt, jolloin niitä ei oteta mukaan käsittely- eikä kontrolliryhmään. Esimerkiksi ensimmäisen kokeilun tapauksessa kaikki yhdeksän suurinta kuntaa olivat mukana kokeilussa, jolloin osa kunnista ei välttämättä täytä ehtoa, jolloin näitä kuntia ei voida ottaa mukaan käsittelyryhmään. Tässä menetetään tietoa, mutta näin on tehtävä, jotta on mahdollista saada uskottavampi asetelma.

Teimme kaltaistamisen molempien kokeilujen tapauksessa käsittelyä edeltävällä periodilla eli ensimmäisen kokeilun tapauksessa poikkileikkausaineistolla, jossa havainnot ovat kokeilua edeltävän periodin (heinäkuu 2011 – joulukuu 2011) keskiarvoja. Jälkimmäisen kokeilun tapauksessa helmikuun 2017 – heinäkuun 2017 aineistolla. Taulukoissa (1) ja (2) esitellään koe- ja vertailuryhmän keskiarvot ennen kaltaistamista ja sen jälkeen ensimmäisen kokeilun osalta, tässä järjestyksessä. Jäljempänä taulukoissa (3) ja (4) esitellään jälkimmäisen kokeilun kuntien osalta tilanne ennen kaltaistamista ja sen jälkeen.

Kuten nähdään, ennen kaltaistamista kokeilukunnat poikkesivat kaikkien taustamuuttujien osalta kokeiluun osallistumattomista kunnista. Asetelma ei siis näytä lähellekään satunnaiskokeen kaltaiselta, sillä satunnaisessa koeasetelmassa koe- ja käsittelyryhmät eroaisivat ainoastaan luottamustason verran eli joka 20. tapauksessa. Taulukossa 2 esitellään tulokset kaltaistamisen jälkeen, ja siitä nähdään, että kaltaistamisella saadaan luotua kontrolliryhmä, joka ei eroa tilastollisesti merkitsevästi käsittelyryhmästä, sillä periodilla, jolla kaltaistaminen on tehty.

Taulukko 1: Kuntakokeilu 2012–2015, taustamuuttajat ennen kaltaistamista

	Koeryhmä, keskiarvo	Verrokkiryhmä, keskiarvo	Erotus, K – V	Keskivirhe	t-testin arvo	p-arvo
Työikäinen väestö	40397,30	6741,96	33655***	4793	7,02	< 0,001
Työttömyysaste	0,0675	0,0533	0,014***	0,003	4,98	< 0,001
Työttömien määrä	2680,46	361,15	2319,3***	281,5	8,24	< 0,001
Toimenpiteet suhteessa työikäiseen väestöön	0,0297	0,0225	0,0071***	0,0015	4,78	< 0,001
Toimenpiteet	1098,5	159,07	939,4***	112,70	8,38	< 0,001
N	249	59				

Taulukko 2: Kuntakokeilu 2012–2015, taustamuuttajat kaltaistamisen jälkeen

	Koeryhmä, keskiarvo	Verrokkiryhmä, keskiarvo	Erotus, K – V	Keskivirhe	t-testin arvo	p-arvo
Työikäinen väestö	20364,02	11663,7	8700*	4748	1,83	0,07
Työttömyysaste	0,066	0,067	-0,0013	0,003	-0,41	0,68
Työttömien määrä	1288,04	744,0	544,0**	253,3	2,1	0,03
Toimenpiteet suhteessa työikäiseen väestöön	0,0297	0,0304	-0,0007	0,0018	-0,37	0,71
Toimenpiteet	566,48	345,16	221,32**	107,86	2,05	0,042
N	51	51				

Teemme kaltaistamisen jälkimmäisen kokeilun tapauksessa samalla tavalla kuin edellä on tehty ensimmäiselle kokeilulle. Kaltaistaminen on siis tehty niin, että jokaiselle jälkimmäiseen kokeiluun osallistuneelle kunnalle on haettu yksi vastinpari. Lisäksi sekä kokeiluun osallistumattomia että siihen osallistuneita kuntia on poistettu, jos oletus yhteisestä tuesta ei ole täyttynyt. Taulukoissa 3 ja 4 esitellään taustamuuttujat ennen kaltaistamista ja sen jälkeen samalla tavalla kuin aiemmin ensimmäisen kokeilun tapauksessa.

Taulukko 3: Kuntakokeilu 2017–2018, taustamuuttujat ennen kaltaistamista

	Koeryhmä, keskiarvo	Verrokkiryhmä, keskiarvo	Erotus, K – V	Keskivirhe	t- testin arvo	p-arvo
Työikäinen väestö	33088	11666	21421,74***	8052	2,66	0,008
Työttömyysaste	0,075	0,067	0,0077*	0,0045	1,69	0,093
Työttömien määrä	2801,98	871,97	1930***	649,5	2,97	0,0032
Toimenpiteet suhteessa työikäiseen väestöön	0,0283	0,0248	0,0035	0,002	1,56	0,12
Toimenpiteet	997,28	330,36	666,9***	225,2	2,96	0,003
N	288	23				

Taulukko 4: Kuntakokeilu 2017–2018, taustamuuttajat kaltaistamisen jälkeen

	Koeryhmä, keskiarvo	Verrokkiryhmä, keskiarvo	Erotus, K – V	Keskivirhe	t-testin arvo	p-arvo
Työikäinen väestö	26371,1	15176,0	11195,14	9029	1,24	0,22
Työttömyysaste	0,0738	0,0787	-0,0049	0,0046	-1,06	0,30
Työttömien määrä	2151,33	1261,3	890,03	828,3	1,08	0,29
Toimenpiteet suhteessa työikäiseen väestöön	0,0280	0,028	-0,0009	0,0027	-0,34	0,74
Toimenpiteet	760,95	527,20	233,75	296,4	0,79	0,43
N	22	22				

LIITE A.2 – Erot eroissa -analyysi

Määritellään jokaiselle periodille k käsittelyindikaattorit D_{it}^k , joille $D_{it}^k = 1$, jos $t = k$ ja kunta i on kokeilukunta. Muulloin $D_{it}^k = 0$. Kun Y_{it} on selitettävä muuttuja (esimerkiksi työttömyysaste), regressiomalli on muotoa:

$$Y_{it} = \alpha + \gamma_i + \lambda_t + \sum_k \beta_k D_{it}^k + \varepsilon_{it}$$

Kuuden kuukauden event study suoritetaan niin, että ensin yhdistetään kuukausitason havainnot kuuden kuukauden osiin, jolloin uusi havainto on kuuden kuukauden keskiarvo. Käsittely alkaa ensimmäisen kokeilun tapauksessa syyskuussa 2012. Esimerkiksi valitsimme ensimmäisen kokeilun tapauksessa tarkasteluperiodiksi yhdeksän vuotta, jolloin tarkastellaan yhteensä 18 periodia, alkaen periodista $k = -7$ ja päättyen periodiin $k = 10$. Regressiomalliin tulee siis 18 käsittelyindikaattoria, jolloin ensimmäisen kokeilun event study -asetelmassa käytetty regressiomalli on muotoa:

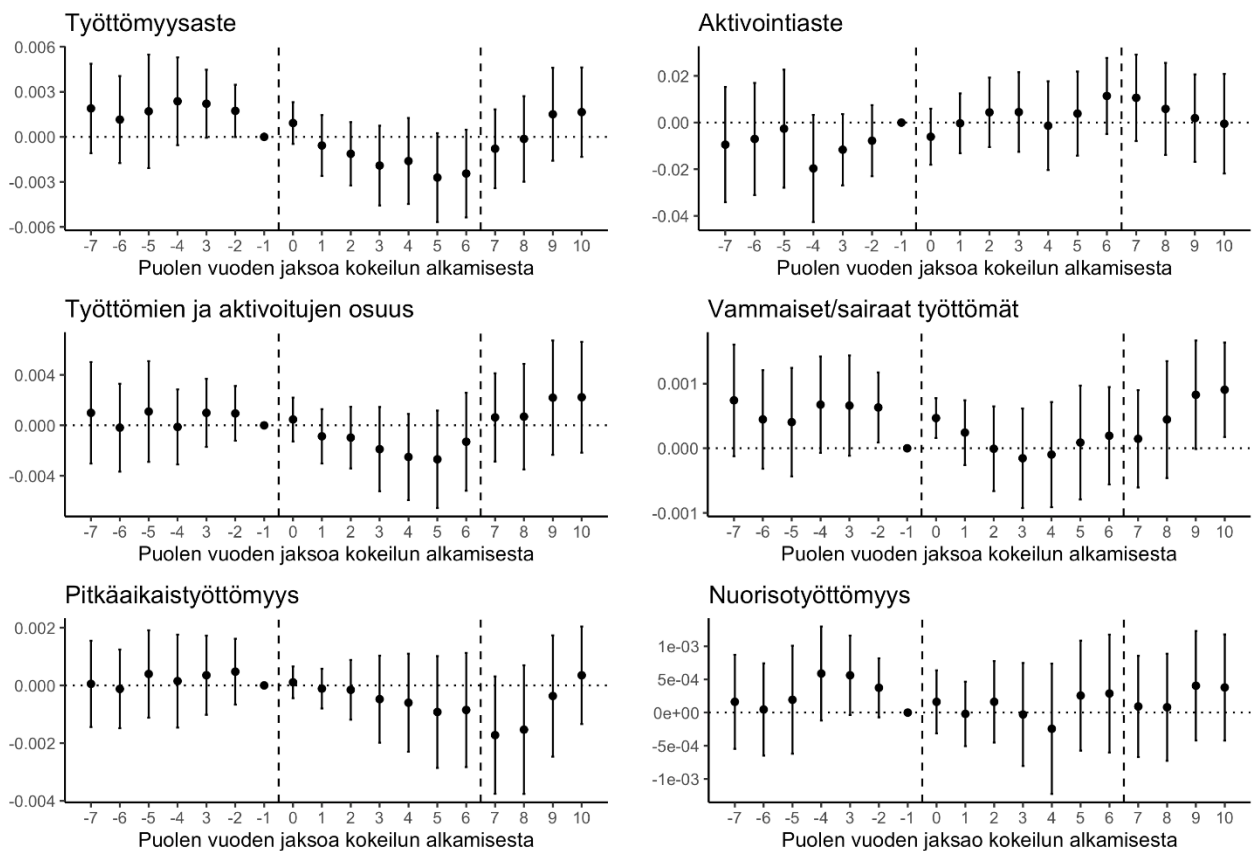
$$Y_{it} = \alpha + \gamma_i + \lambda_t + \sum_{k=-7}^{10} \beta_k D_{it}^k + \varepsilon_{it}$$

Mallista jätetään yksi, käsittelyn alkamista edeltävän periodin käsittelyindikaattori pois. Saadut kertoimet ovat siis suhteessa periodin $k = -1$ eli kokeilun alkamista edeltäneen puolivuotisen tilanteeseen. Toisen kokeilun tapauksessa käsittely alkaa elokuussa 2017. Valitsimme tarkasteluperiodiksi seitsemän vuoden ajanjakson. Regressiomalliin tulee siis 13 käsittelyindikaattoria, sillä jälleen yksi kerroin (kokeilun alkamista edeltävän puolivuotisen kerroin eli $k = 11$) normeerataan nolllaksi. Malli on siis muotoa:

$$Y_{it} = \alpha + \gamma_i + \lambda_t + \sum_{k=-11}^5 \beta_k D_{it}^k + \varepsilon_{it}$$

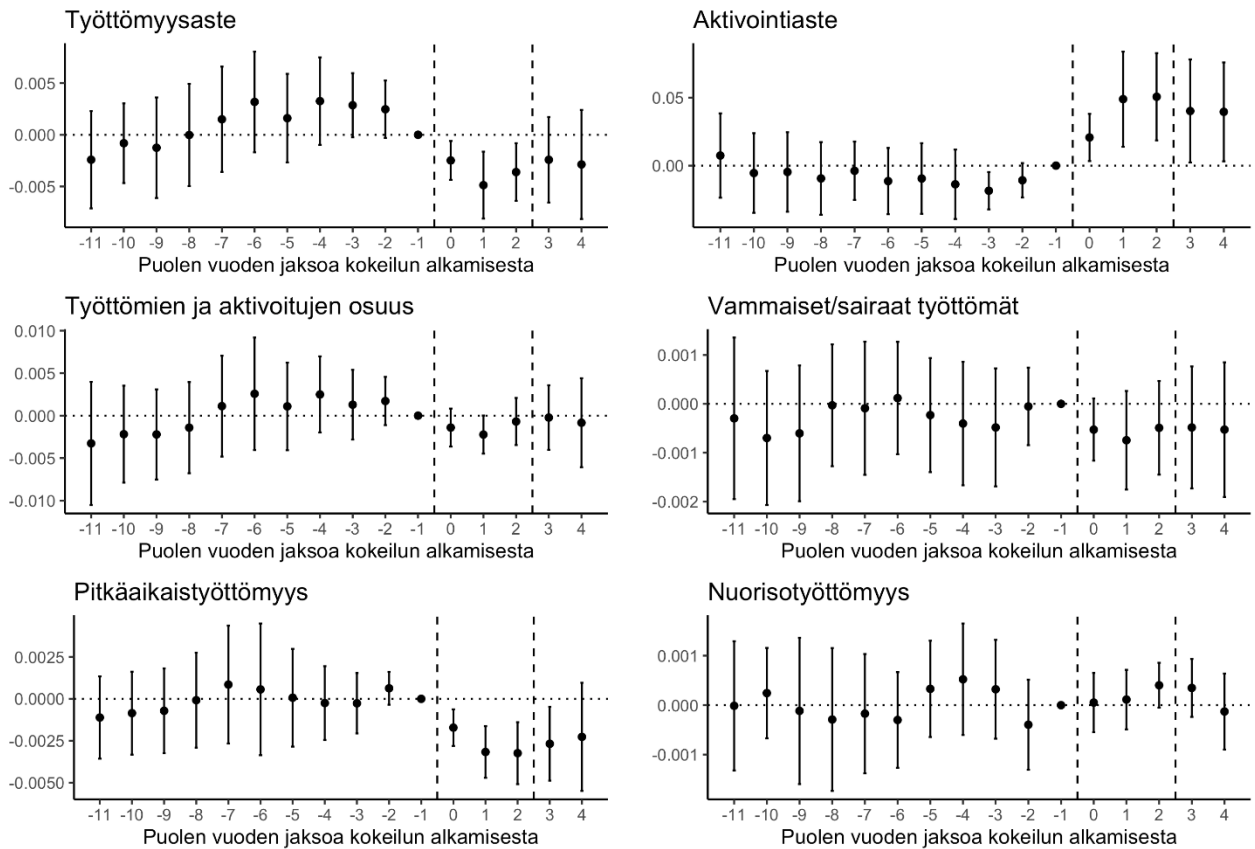
LIITE A.3 – Keskivirheiden klusterointi

Paneeliregressioita tehtäessä keskivirheiden klusterointia pidetään usein tärkeänä. Tässäkin tapauksessa klusterointia tulisi miettiä, sillä molemmissa kokeiluissa käsittelykunnat kuuluvat kuntaryppäisiin: ensimmäisessä kokeilussa ryppäitä on 15 kappaletta ja toisessa 5 kappaletta. Klusterointi voidaan suorittaa esimerkiksi kuntatasolla tai seutukunnan tasolla. Tuloksissa hämmästyttävää on, että joidenkin regressiokertoimien keskivirheet pienenevät, kun ne klusteroidaan seutukunnan tasolla. Yleensä klusterointi kasvattaa keskivirheitä, joten keskivirheiden pieneneminen on epätavallista. Syy klusteroitujen keskivirheiden outoon käyttäytymiseen saattaa olla aggregaattitason aineisto tai negatiivinen korrelaatio klusterin sisällä. Kuviossa 1 on ensimmäisen kokeilun ja kuviossa 2 on jälkimmäisen kokeilun event study, jossa keskivirheet on klusteroitu seutukunnan tasolla.



Kuvio 3: Event study (2012–2015), keskivirheet klusteroitu seutukunnan tasolla

Selite: Muut muuttujat suhteessa työikäiseen väestöön, aktivointiaste suhteessa työttömien ja aktivoitujen määrään. Kuvioden kertoimet ovat käsittelyindikaattorien kertoimia regressiomallista, jossa on kiinteät selittäjät kunnalle ja aikaperiodille (kuuden kuukauden jaksot). Keskivirheet on klusteroitu seutukunnan tasolla. Kokeilu-aika (9/2012–12/2015) merkitty katkoviivoilla. Koeryhmässä 51 kuntaa, kontrolliryhmässä 51 kuntaa.

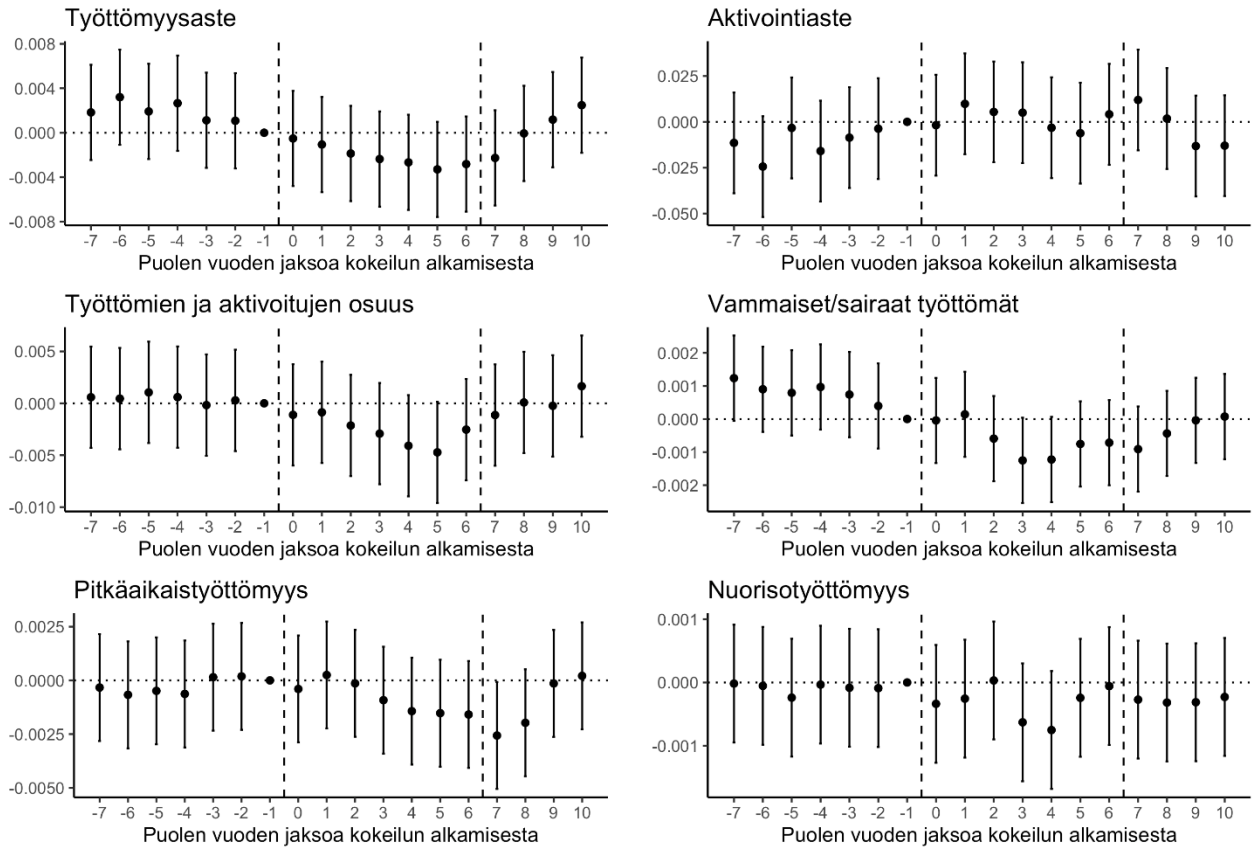


Kuvio 4: Event study (2017–2018), keskirviheet klusteroitu seutukunnan tasolla

Selite: Muut muuttujat suhteessa työikäiseen väestöön, aktivointiaste suhteessa työttömien ja aktivoitujen määrään. Kuvioiden kertoimet ovat käsittelyindikaattorien kertoimia regressiomallista, jossa on kiinteät selittäjät kunnalle ja aikaperiodille (kuuden kuukauden jaksot). Keskirviheet on klusteroitu seutukunnan tasolla. Kokeilu-aika (8/2017–12/2018) merkitty katkoviivoilla. Koeryhmässä 22 kuntaa, kontrolliryhmässä 22 kuntaa.

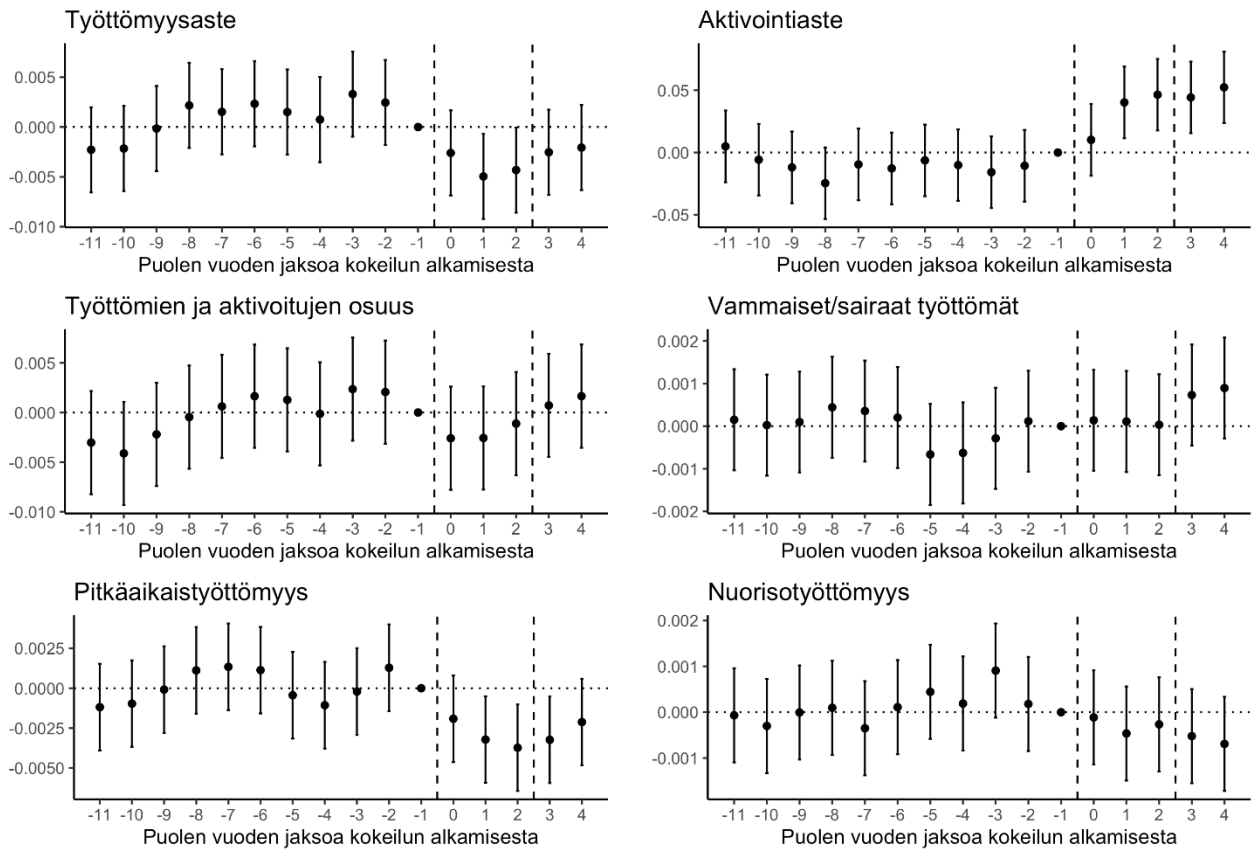
Kuten kuviosta 2 nähdään, asetelma näyttää jonkin verran huonommalta jälkimmäisen kokeilun työttömyysasteen kohdalla, sillä joistakin kokeilua edeltävistä kertoimista tulee merkitseviä. Tämä on hieman ongelmallista erotusten erotus -menetelmän oletusten kannalta. Muiden vastemuuttujien, kuten pitkäaikaistyöttömyyden tai aktivointiasteen kohdalla näin ei käy, vaan asetelma näyttää yhtä hyvältä kuin ilman keskirviheiden klusterointia.

LIITE A.4 – Hakeneiden kuntien tarkastelu



Kuvio 5: Event study (2012–2015), hakeneet kunnat

Selite: Muut muuttujat suhteessa työikäiseen väestöön, aktivointiaste suhteessa työttömien ja aktivoitujen määrään. Kuvioden kertoimet ovat käsittelyindikaattorien kertoimia regressiomallista, jossa on kiinteät selittäjät kunnalle ja aikaperiodille (kuuden kuukauden jaksot). Kokeiluaika (9/2012–12/2015) merkitty katkoviivoilla. Koeryhmässä 34 kuntaa, kontrolliryhmässä 34 kuntaa. Koe- ja kontrollikuntien vähäisempi määrä johtuu siitä, että hakeneiden kuntien tarkastelussa kuntia tippuu kaltaistamisessa enemmän.



Kuvio 6: Event study (2017–2018), hakeneet kunnat

Selite: Muut muuttujat suhteessa työikäiseen väestöön, aktivointiaste suhteessa työttömien ja aktivoitujen määrään. Kuvioiden kertoimet ovat käsittelyindikaattorien kertoimia regressiomallista, jossa on kiinteät selittäjät kunnalle ja aikaperiodille (kuuden kuukauden jaksot). Kokeiluaika (8/2017–12/2018) merkitty katkoviivoilla. Koeryhmässä 22 kuntaa, kontrolliryhmässä 22 kuntaa.

LIITE A.5 – Koe- ja kontrollikunnat

Koekunnat (2012–2015) otoksessa, 51 kpl:

Akaa, Espoo, Hämeenlinna, Hamina, Hartola, Hattula, Heinola, Hollola, Imatra, Jämsä, Janakkala, Kajaani, Kemi, Keuruu, Kokkola, Kotka, Kuhmo, Kuopio, Lappeenranta, Lemi, Leppävirta, Lieksa, Luumäki, Merikarvia, Mikkeli, Muurame, Orimattila, Outokumpu, Parikkala, Pomarkku, Pori, Pyhäjoki, Raahe, Ranua, Rautalamppi, Rautjärvi, Rovaniemi, Ruokolahti, Savitaipale, Savonlinna, Seinäjoki, Siikajoki, Siilinjärvi, Suonenjoki, Sysmä, Taipalsaari, Tuusniemi, Ulvila, Urjala, Valkeakoski, Varkaus

Kontrollikunnat (2012–2015), 51 kpl:

Äänekoski, Enontekiö, Forssa, Haapavesi, Huittinen, Hyvinkää, Ii, Iisalmi, Jämijärvi, Kangasniemi, Kankaanpää, Kannus, Kärsämäki, Kaskinen, Kemijärvi, Kempele, Kinnula, Kitee, Kittilä, Konnevesi, Kontiolahti, Kouvola, Kuusamo, Kyyjärvi, Laitila, Loimaa, Loviisa, Mänttä-Vilppula, Muhos, Naantali, Nokia, Nurmes, Pieksämäki, Porvoo, Pyhäntä, Rantasalmi, Rauma, Riihimäki, Ristijärvi, Saarijärvi, Salla, Salo, Sastamala, Suomussalmi, Tohmajärvi, Toivakka, Tornio, Utsjoki, Vaala, Vaasa, Veteli

Kaltaistamisessa tippuneet koekunnat (2012–2015), 8 kpl:

Helsinki, Joensuu, Jyväskylä, Lahti, Oulu, Tampere, Turku, Vantaa

Koekunnat (2017–2018) otoksessa, 22 kpl:

Kangasala, Kemi, Kemijärvi, Kuopio, Lempäälä, Naantali, Nokia, Orivesi, Paimio, Pirkkala, Pori, Punkalaidun, Raisio, Rovaniemi, Sastamala, Siilinjärvi, Sodankylä, Tornio, Turku, Tuusniemi, Vesilähti, Ylöjärvi

Kontrollikunnat (2017–2018) otoksessa, 22 kpl:

Haapavesi, Hamina, Juuka, Jyväskylä, Kiuruvesi, Kokkola, Lieksa, Loviisa, Maalahti, Merijärvi, Muonio, Parikkala, Pello, Pyhtää, Raahe, Ristijärvi, Salo, Simo, Vaasa, Vesanto, Ylitornio, Ylivieska

Kaltaistamisessa tippuneet koekunnat (2017–2018), 1 kpl:

Tampere