

Mehtonen Mikko

Samankaltaiset kunnat

SAKU

Menetelmäkuvaus 11.11.2021



© Suomen Kuntaliitto ry
Helsinki 2021

Kuntaliitto
Toinen linja 14
PL 200, 00101 Helsinki
Puhelin 09 7711
www.kuntaliitto.fi

Sisältö

1	Menetelmäkuvaus	4
1.1	Teema1: kunnan kustantama varhaiskasvatus	6

1 Menetelmäkuvaus

SAKU samankaltaiset kunnat -hankkeessa etsitään kuntien kustannuseroja selittäviä tekijöitä. Löydettyjen tekijöiden perusteella jokaiselle kunnalle valitaan mahdollisimman samankaltaiset vertailukunnat, joihin kunta voi verrata omaa toimintaansa.

Tutkimus toteutetaan teemoittain. Alkuvaiheeseen valittuja teemoja ovat: varhaiskasvatus, esiopetus, perusopetus sekä muuttoliike. Opetuspalveluiden teemoissa selvitetään, mitkä tekijät selittävät kustannuseroja. Muuttoliike poikkeaa edellä mainituista teemoista siten, että kustannuserojen sijasta tutkitaan, mitkä tekijät selittävät kuntien välistä muuttoliikettä.

Hanke aloitettiin haastattelemalla kunkin teeman, esimerkiksi varhaiskasvatuksen, asiantuntijoita. Haastattelujen avulla kerättiin mahdollisimman laaja lista tekijöistä, joilla saattaa olla yhteys kustannuksiin. Seuraavassa vaiheessa selvitettiin, mitkä listauksen tekijöistä on muutettavissa numeeriseen muotoon eli muuttujiksi ja joihin löytyy numeerista tilastoaineistoa. Kunnilta ei kerätty aineistoa, vaan hankkeessa käytettiin ainoastaan jo kerättyä tilastotietoa. Osa SAKU-malliin ehdolla olevista tekijöistä karsiutui sen vuoksi, ettei tilastoaineistoa ollut saatavilla. Tältä osin tutkimusta voisi kehittää tulevaisuudessa.

Muuttujiksi asti määritellyt tekijät jaoteltiin olosuhde- sekä päätösperäisiin muuttujiin. Olosuhdetekijät ovat sellaisia, joihin kunta ei voi itse merkittävästi vaikuttaa. Päätösperäiset tekijät taas ovat sellaisia, joihin kunta voi vaikuttaa. Tilastollisiin malleihin valittiin sekä olosuhde- että päätösperäisiä muuttujia.

SAKU-hankkeen tilastollisen tutkimuksen vaiheessa selvitettiin regressioanalyysin avulla edellä valittujen muuttujien selitysvoimaa. Regressiomallin avulla muuttujalistasta valittiin ne muuttujat, jotka selittivät parhaiten kuntien välisiä kustannuseroja. Regressiomalleja testattiin useammilla eri muuttujakombinaatioilla muun muassa siten, että keskenään vahvasti korreloivista muuttujista valittiin malliin vain toinen.

Regressiomalleissa hyödynnettiin ajallista ja alueellista vaihtelua. Etenkin alueellinen vaihtelu selittää eroja, koska se on lähtökohtaisesti suurempaa.

Aineisto koostuu kuntakohtaisesta aineistosta vuosilta 2015–2019, lukuun ottamatta varhaiskasvatusta, jonka kuntakohtainen aineisto koostuu vuosien 2016–2019 tilastotiedoista.

Estimoiteja kokeiltiin myös kuntien kiinteillä vaikutuksilla. Ne eivät kuitenkaan vaikuttaneet tuloksiin (testattu järjestämistapaestimaatilla), eli nykyiset tulokset ovat vahvat. Kiinteät vaikutukset poistavat alueellisen vaihtelun eli kuntien väliset erot, jolloin päädytään seuraamaan vain ajallista vaikutusta. Kiinteiden vaikutusten estimointia ei käytetty varsinaisessa SAKU-mallissa.

Varsinaiseen SAKU-malliin valittiin regressioanalyysin perusteella muuttujat joissa:

- regressiokerroin oli tilastollisesti merkitsevä ($p < 0.05$),
- tilastollinen merkitsevyys säilyi useammassa regressioajoissa (eri muuttujakombinaatioilla),
- regressiokerroin oli tarpeeksi suuri,
- muuttuja ei korreloi muiden malliin valittujen muuttujien kanssa (suora korrelaatio maksimissaan 0,8).

SAKU-malliin valittiin ne muuttujat, joissa edellä mainitut ehdot täyttyivät. Samankaltaisia kuntia etsittäessä, malliin valittuja muuttujia painotetaan regressiotulosten perusteella. Painoarvot laskettiin kertomalla muuttujan regressiokerroin kvartiilivälin pituudella. Muuttujakohtainen tulos suhteutettiin kaikkien malliin valittujen muuttujien yhteenlaskettuun tulokseen ja näin saatiin selville kunkin muuttujan prosentuaalinen painoarvo. Edellä mainittu laskentatoimenpide toteutettiin sen vuoksi, että muuttujat eivät ole samalla asteikolla.

Muuttujakohtaisten painoarvojen laskennassa regressiokerroin olisi voitu kertoa myös keskihajonnalla. Keskihajonta olisi kuitenkin toiminut huonosti joidenkin, jakaumaltaan vinoutuneiden muuttujien, kuten väkiluvun, osalta. Tämän vuoksi painotuslaskennassa päädyttiin käyttämään kvartiilivälin pituutta.

Eräät muuttujat karsiutuivat SAKU-mallista sen vuoksi, että niiden painoarvo jäi edellä mainitulla laskentatavalla hyvin pieneksi. Tämä johtui regressiokertoimen alhaisesta arvosta. Esimerkiksi varhaiskasvatuksen mallissa väkiluku selitti kustannuseroja tilastollisesti merkitsevästi, mutta niin vähän, että väkilukua ei käytetä samankaltaisten kuntien etsimisessä.

Regressiomalli toimii hyvin silloin, kun siinä on mukana kaikki selitettävään ilmiöön vaikuttavat selittävät tekijät ja niitä mittaavat muuttujat. Puuttuvan muuttujan harha syntyy silloin, kun mallista puuttuu sellainen muuttuja, jolla olisi selitysvoimaa ja joka korreloi yhden tai useamman mallissa olevan selittävän muuttujan kanssa.

1.1 Teema 1: kunnan kustantama varhaiskasvatus

Teemassa etsitään kuntien varhaiskasvatuksen kustannuseroja selittäviä tekijöitä. Kunnan kustantamalla varhaiskasvatuksella tarkoitetaan kunnan omaa perhepäivä- ja päiväkotihoitoa, palvelusetelillä järjestettyä toimintaa, kunnan maksamaa yksityistä varhaiskasvatusta, kuntalisiä sekä muuta varhaiskasvatusta. Eli kustannukset muodostuvat varhaiskasvatuksen kokonaiskustannuksista.

Taloustietojen lähteenä käytetään Tilastokeskuksen taloustilastoa. Kustannuksina käytetään käyttökustannuksia eli niihin sisältyivät toimintakulut, poistot ja arvonalentumiset sekä sisäiset erät ja vyörytykset. Kustannuksista ei vähennetty asiakasmaksuja, mutta niistä vähennettiin muille ulosmyyty palvelu eli kustannusten pitäisi kuvata mahdollisimman hyvin oman kunnan asukkaille tarjottujen varhaiskasvatuspalveluiden kustannuksia. Näitä kustannustietoja on mielekästä käyttää, kun olosuhde- ja päätösmuuttujat kuvaavat myös kyseisen kunnan tilannetta.

Selitettävä muuttuja: varhaiskasvatuksen käyttökulut / kunnan 1-6 vuotias

Varhaiskasvatuksen viimeisimmän regressiomallin selitysaste on 68 %. Ilman mekaanisia kustannustekijöitä selitysaste on 63 %.

Selitettävän muuttujan laskennassa kunnan käyttökulut suhteutettiin aluksi suoraan varhaiskasvatuksen asiakasmäärään. Tämä laskentatapa ei kuitenkaan tuottanut regressioajoissa merkittäviä tuloksia. Mallien selitysaste nousi merkittävästi kun kustannukset suhteutettiin palvelua käyttävään ikäluokkaan. Todennäköinen syy tälle on se, että julkisista lähteistä saadut asiakastiedot eivät olleet vertailukelpoisia kustannustietojen kanssa. Mallia voitaisiin jatkokehittää, jos selitettävä muuttuja saataisiin jatkossa laskettua luotettavasti muodossa euroa per asiakas.

Asiantuntijahaastatteluissa tunnistettiin yli 40 varhaiskasvatuksen kustannuksiin mahdollisesti vaikuttavaa tekijää. Kustannuserojen selittäjiksi näistä päästiin testaamaan yli 30 muuttujaa:

Kunnan kustantamaan varhaiskasvatukseen osallistuneiden (osallistumisaste*1-6 -vuotias väestö) osuus väestöstä, %	***
---	-----

Kunnan kustantamaan varhaiskasvatukseen osallistuneiden (osallistumisaste*1-6 -vuotias väestö) vuosimuutos, %	***
Kunnan kustantamaan varhaiskasvatukseen osallistuneet 31.12., % 1 - 6-vuotiaista	***
Ansiotulot per asukas, eur./as.	***
Kelan tilastoissa positiivista kuntalisää maksaneet kunnat = 1, muut kunnat = 0	***
Koko- tai osa-aikaiseen päiväkotitai perhepäivähoitotoimintaan osallistuneiden lasten lukumäärän suhde lukuun, johon on laskettu edellisen lisäksi myös yksityisen hoidon tuella hoidetut lapset ja palveluseteliasiakkaat	***
Varhaiskasvatustoiminnan vuokratkustannukset (sisäiset+ulkoiset) per lapsi (= 1-6 v. väestö)	***
Taloudellinen huoltosuhte, %	***
Verotulot per asukas, eur./as.	***
Valtionosuuden osuus verorahoituksesta (verotulot + valtionosuudet), %	***
Väkiluku 31.12, lkm	***
Taajama-aste, %	**
Maanteiden yhteispituus, km	**
Keskiasteen koulutuksen saaneet, % 20 vuotta täyttäneistä	**
Korkea-asteen koulutuksen saaneet, % 20 vuotta täyttäneistä	**
Perusasteen jälkeen suoritetun korkeimman koulutuksen keskimääräinen pituus henkeä kohti	**
Kunnan järjestämään kokoaikaiseen varhaiskasvatukseen osallistuneiden 1-2 -vuotiaiden osuus vastaavalla tavalla luokitelluista 1-5 -vuotiaista, %	*
Kehyskunta = 1, muu kunta = 0	*
Työttömyysaste, %	*
Keski-ikä, vuotta	ei selitystasetta
Väkiluvun vuosimuutos, %	ei selitystasetta
Maapinta-ala, km ²	ei selitystasetta
Saaristokunta = 1, muu kunta = 0	ei selitystasetta
Saaristo-osakunta = 1, muu kunta = 0	ei selitystasetta
Liikennekäyttöisten autojen lukumäärä per asukas	ei selitystasetta
Vasemmistoenemmistöinen (SDP/Vasemmistoliitto) valtuusto = 1, muu valtuusto = 0	ei selitystasetta
Toimeentulotukea lyhytaikaisesti saaneet yksinhuoltajakotitaloudet, % toimeentulotukea saaneista kotitalouksista	ei selitystasetta
Huostassa vuoden aikana olleet 3 - 6-vuotiaat, % vastaavanikäisistä	ei selitystasetta

Koko- tai osa-aikaiseen päiväkotitoimintaan osallistuneiden lasten lukumäärän (oma toiminta) suhde lukuun johon on laskettu edellisen lisäksi myös koko- tai osa-aikaiseen perhepäivähoitoon osallistuneiden lasten lukumäärä (oma toiminta), %	ei selitysastetta
Pidennetyn oppivelvollisuuden oppilaat, % perusopetusoppilaista	ei selitysastetta
Ruotsinkielisten osuus väestöstä, %	ei selitysastetta
Vieraskielisten osuus väestöstä, %	ei selitysastetta
Toimialarakenne	ei selitysastetta

*** p<0.01, ** p<0.05, * p<0.1

Ikärakennetta kokeiltiin aluksi eri ikäryhmien väestöosuuksina. Kunnan ikärakennetta kuitenkin yksinkertaistettiin, jolloin päädyttiin kunnan keski-ikää kuvaavaan muuttajaan. Keski-ikä ei kuitenkaan havaittu selittävän kustannuseroja.

Edellä mainituista muuttujista SAKU-malliin valittiin seuraavat muuttajat:

- Taloudellinen huoltosuhte (Painoarvo 24 %)
- Maanteiden pituus: (Painoarvo 4 %)
- Kuntalisä: (Painoarvo 5 %)
- Osallistumisaste: (Painoarvo 27 %)
- Omatoiminnan osuus: (Painoarvo 7 %)
- Varhaiskasvatuksen asiakasmäärän osuus väestöstä: (Painoarvo 19 %)
- Ansiotulot asukasta kohden: (Painoarvo 14 %)

Kaikki edellä mainitut muuttajat selittivät kustannuseroja tilastollisesti merkitsevästi ja merkitsevyys säilyi hyvänä eri muuttujakombinaatioilla. Prosentit kuvaavat muuttujalle laskettua painoarvoa.

Regressioanalyysi ei kerro syytä siihen, miksi jokin asia selittää kuntien välisiä kustannuseroja. Muuttujien yhteyttä kustannuksiin voidaan kuitenkin arvioida toimialatuntemuksen avulla.

Taloudellinen huoltosuhte ja osallistumisaste saivat laskelmissa korkeimman painoarvon ja ne siis selittävät mallissa testatuista muuttujista parhaiten kustannuseroja. Molempien muuttujien tulkinta tuntuu hyvin intuitiiviselta. Kun selitettävän muuttujan kustannukset on suhteutettu ikäluokkaan, liittyy ainakin osallistumisaste suoraan palvelun käyttöön ja siten kustannuksiin. Taloudellinen huoltosuhte liittyy myös ainakin palveluiden käyttöön, mutta mahdollisesti myös johonkin muuhun selittävään tekijään. Molemmat muuttajat pidettiin mallissa, koska ne eivät korreloineet vahvasti keskenään (0.18).

Maanteiden pituus näytti selittävän kustannuseroja ja sen voidaan olettaa liittyvän palveluverkkokysymyksiin. Pitkä maantieverkosto kuvaa laajalle levittäytynyttä asukasverkostoa paremmin kuin kunnan pinta-ala, joka ei saanut tilastollista merkitsevyyttä. Maantiet mahdollistavat laajasti asumisen ja maanteiden ympärille muodostuu taajamia. Laajalle levinnyttä palveluverkkoa on vaikeampi järjestää tehokkaasti, mikä näkyy kustannuksissa.

Kuntalisän vaikutus kustannuksiin tuntuu lähes itsestään selvältä tulokselta. Kuntalisän merkitys on tässä kohden kuitenkin monitulkintaisempi. Kuntalisän maksamisella tavoitellaan yleensä alhaisempaa osallistumisastetta, jolloin kuntalisän aiheuttamien lisäkustannusten pitäisi kompensoitu alhaisempina oman toiminnan kustannuksina. Kyseiset muuttajat eivät kuitenkaan korreloineet keskenään (0.11) ja myös kuntalisä sekä osallistumisaste päätettiin pitää mallissa. Kuntalisää maksaneilla kunnilla kustannukset olivat lasta kohden keskimäärin noin 400 euroa korkeammat.

Oman toiminnan osuus selittää myös kustannuseroja ja muuttuja valittiin osaksi SAKU-mallia. Oman toiminnan osuus sai regressiossa negatiivisen kertoimen eli ulkoistaminen näyttäisi korottavan kustannuksia. Tuloksista ei voida kuitenkaan suoraan päätellä ulkoistamisen olevan kalliimpaa. Yksi selitys ilmiölle voisi olla se, että ulkoistamisella on nouseva rajakustannus. Aluksi ulkoistetaan yksinkertaista päiväkotitoimintaa ja kuntiin jää vaikeammat tapaukset, kuten erityisopetus ja muu kalliimpi päivähoito. Kun ulkoistetaan enemmän, alkaa vaativampaa toimintaa siirtymään myös ulkoistamisen piiriin. SAKU-mallin kannalta tärkeä havainto on kuitenkin se, että oman toiminnan osuudella on yhteys kustannuksiin ja samankaltaisia kuntia valitteassa muuttuja on syytä pitää mallissa.

Varhaiskasvatuksen skaalaetuja testattiin väkimäärään suhteutetulla asiakasosuudella sekä asiakasosuuden muutoksella. Skaalaetuja ei voitu testata absoluuttisella asiakasmäärällä, koska se korreloisi liikaa regressioajoissa olevan väkiluvun kanssa. Sekä asiakasosuus että asiakasosuuden muutos saivat negatiivisen regressiokertoimen. Tulosten perusteella varhaiskasvatuksessa näyttäisi siten olevan skaalaetuja, mikä selittyy sillä, että kun verrataan väestöltään, osallistumisasteeltaan ja palveluverkoltaan (sekä muut selittävät muuttajat) samankaltaisia kuntia, näkyy asiakasosuuden kasvu laskevina kustannuksina. SAKU-malliin valittiin näistä kahdesta muuttujasta asiakasmäärän osuus väestöstä, koska se sai suuremman painoarvon isomman regressiokertoimen ansiosta.

Kunnan asukkaiden ansiotulot asukasta kohden on toistaiseksi viimeisin SAKU-malliin valittu muuttuja. Ansiotulojen hyvä selitysaste ja positiivinen regressiokerroin antaa viitteitä siitä, että hyvän taloudellisen tilanteen kunnissa

varhaiskasvatuksen palveluihin panostetaan, joka näkyy korkeampina kustannuksina. Korkean tulotason kunnissa ei useinkaan ole maan korkein kunnallisten sote-palveluiden kysyntä, joka taas antaa kunnalle taloudellista liikkumavaraa panostaa esimerkiksi opetuksen palveluihin. Ansiotulojen pitämistä SAKU-mallissa tulisi kuitenkin pohtia jatkokehittämisen yhteydessä, koska kyseinen muuttuja korreloi jossain määrin taloudellisen huoltosuhteen kanssa (-0.64).