

VATT TUTKIMUKSET 195

Kuntien palvelutuotannon menot ja valtionosuuksien määräytymistekijät

Teemu Lyytikäinen
Markku Siikanen

VATT TUTKIMUKSET 195

VATT TUTKIMUKSET 195

Kuntien palvelutuotannon menot ja valtionosuuksien määräytymistekijät

Teemu Lyytikäinen

Markku Siikanen

Teemu Lyytikäinen, Valtion taloudellinen tutkimuskeskus,
teemu.lyytikainen@vatt.fi

Markku Siikanen, Valtion taloudellinen tutkimuskeskus,
markku.siikanen@vatt.fi

ISBN 978-952-274-298-8 (PDF)

ISSN 1795-3340 (PDF)

URN:ISBN:978-952-274-298-8

Valtion taloudellinen tutkimuskeskus
VATT Institute for Economic Research
Arkadiankatu 7, 00100 Helsinki, Finland

Helsinki, syyskuu 2024

Esipuhe

Tämä tutkimusraportti on valtiovarainministeriön (VM) VATT:lta tilaama ja sen tavoitteena on tukea valtionosuusjärjestelmän uudistamista. Haluamme kiittää VM:n Unna Heimbergiä avusta aineiston kokoamisesta sekä VM:n Unna Heimbergiä, Teija Kauhasta, Kati Jussilaa, Markku Nissistä, Arto Sulosta, Jussi Lammassaarta, Lauri Piiraista ja Eero Laesterää hyödyllisistä kommentteista. Lisäksi kiitämme valtionosuusuudistuksen valmisteluryhmän jäseniä hyödyllisestä keskustelusta ja kommentteista ryhmän kokouksissa. Kiitämme myös Alex Kivimäkeä avusta raportin kuvien ja taulukoiden muotoilussa sekä Ramin Izadia esikoulukokeilua koskevan aineiston toimittamisesta.

Helsingissä, 23.8.2024

Teemu Lyytikäinen

johtava tutkija

Markku Siikanen

erikoistutkija

Sisällys

1. Johdanto	1
2. Nykyinen valtionosuusjärjestelmä ja sen huomioiminen tutkimuksessa ...	2
2.1. Nykyinen valtionosuusjärjestelmä	2
2.2. Nykyisten valtionosuuskriteerien huomioiminen tässä tutkimuksessa	5
3. Aineisto ja menetelmät	6
3.1. Tutkimusaineisto	6
3.2. Tutkimusmenetelmä.....	10
4. Kuntien toimintamenojen tarkastelu.....	12
4.1. Kuntien toimintamenojen kuvaileva tarkastelu.....	12
4.2. Kuntien toimintamenojen regressioanalyysi	14
4.3. Regressioanalyysien herkkyystarkastelu.....	17
5. Tulosten hyödyntäminen valtionosuusjärjestelmän uudistamisessa	28
6. Lopuksi.....	29
Lähteet.....	29
Liite A: Muuttujakuvaukset.....	30
Toimintamenot	30
Muut muuttujat	31

1. Johdanto

Kuntien rooli muuttui merkittävästi vuoden 2023 alussa, kun sosiaali- ja terveystalvet sekä pelastustoimi siirtyivät sote-uudistuksessa hyvinvointialueiden vastuulle. Kuntien valtiosuusjärjestelmään tehtiin sote-uudistuksen yhteydessä vain välttämättömät muutokset. Uudistuksen jälkeen kuntien valtiosuusjärjestelmän kehittämistarpeita on perusteltua arvioida suhteessa kuntien jäljelle jääviin tehtäviin ja toimintaympäristön muutoksiin.

Pääministeri Petteri Orpon hallituksen ohjelman mukaan kuntien rahoituksen ja valtiosuusjärjestelmän kokonaisuus uudistetaan vastaamaan kuntien uutta roolia ja sote-uudistuksen voimaantulon sekä TE-uudistuksen jälkeistä tilannetta. Valtiosuus uudistusta valmistelemaan on asetettu ministeriöiden virkahenkilöistä ja muista asiantuntijoista koostuva valmisteluryhmä, jonka työtä tämä tutkimus tukee. Tutkimus on Valtiovarainministeriön tilaama ja rahoittama.

Tämän tutkimuksen tarkoituksena on tutkia kuntien palvelutuotannon menoja selittäviä tekijöitä kuntien peruspalvelujen valtiosuusjärjestelmän uudistamista varten. Tutkimuksessa keskitytään kunnille sote-uudistuksen jälkeen jääneiden varhaiskasvatuksen, esi- ja perusopetuksen, kirjastojen, taiteen perusopetuksen ja yleisen kulttuuritoimen tehtävien menoihin. Tavoitteena on tuottaa arviot kunnan väestön ikärakenteen sekä muiden mahdollisten valtiosuus kriteerien yhteydestä tarkasteltujen palveluiden tarjoamisen ja tuottamisen menoihin.

Vuonna 2025 voimaan tulevasta TE-palvelujen uudistuksesta johtuvia kuntien menoja ja valtiosuusjärjestelmän muutostarvetta ei tarkastella tässä tutkimuksessa. Tutkimuksessa ei myöskään tehdä laskelmia valtiosuusien jakautumisessa tapahtuvista muutoksista, jos valtiosuusjärjestelmän kriteerit ja perushinnat muutettaisiin tutkimuksen tarkastelun mukaisiksi.

Tutkimuksessa käytetään eri lähteistä koottua aineistoa kuntien toimintamenoista ja niitä mahdollisesti selittävistä ominaisuuksista vuosilta 2018–2022. Kuntien ominaisuuksien yhteyttä toimintamenoihin tutkitaan regressiomallilla, jossa sote-uudistuksen jälkeisiä palveluja vastaavia toimintamenoja selitetään mahdollisilla valtiosuus kriteereillä. Tulokset eivät välttämättä anna luotettavaa kuvaa menojen ja kuntien ominaisuuksien välisestä syy-seuraus suhteesta, sillä tutkittavien muuttujien kautta saattaa kanavoitua mallin ulkopuolisten tekijöiden vaikutuksia ja tuloksiin liittyy usein myös suurta tilastollista epävarmuutta varsin pienen otoskoon takia. Regressioanalyysin tuloksia voidaan silti pitää hyvänä lähtökohtana nykyisen valtiosuusjärjestelmän uudistamisessa. Ne tarjoavat objektiivisen vertailukohtan nykyisen valtiosuusjärjestelmän kertoimille.

Tämän tutkimuksen kaltaisia analyyseja on tehty aiemminkin valtiosuusjärjestelmän uudistusten yhteydessä (ks. Harjunen ja Saarimaa, 2012; Lehtonen ym., 2008a ja 2008b sekä Häkkinen ym., 2009). Kuntien menojen ja mahdollisten valtiosuus kriteerien yhteyttä on syytä tarkastella uudelleen, sillä aiempien tutkimusten aineisto ei ole enää ajantasaista eikä sovellu sote-uudistuksen jälkeiseen tilanteeseen.

Raportin luvussa 2 kuvataan nykyinen valtiosuusjärjestelmä. Luvussa 3 esitellään tutkimusaineisto ja tutkimusmenetelmät. Luvussa 4 raportoidaan empiirisen analyysin tulokset. Luvussa 5 käsitellään tulosten hyödyntämistä valtiosuusjärjestelmän uudistamisessa. Luvussa 6 tehdään huomioita vastaavien tutkimusten toteuttamisen edellytyksistä jatkossa.

2. Nykyinen valtionosuusjärjestelmä ja sen huomioiminen tutkimuksessa

2.1. Nykyinen valtionosuusjärjestelmä

Kunnan peruspalvelujen valtionosuus muodostuu kuntien palvelujen kustannuserojen tasauksesta ja laskennallisiin verotuloihin perustuvasta valtionosuuden tasauksesta. Tässä tutkimuksessa keskitytään kustannuserojen tasaukseen, joka perustuu kunnan asukkaiden palvelutarpeeseen ja olosuhdetekijöihin. Palvelutarvetta ja tuotantokustannuksiin vaikuttavia olosuhdetekijöitä pyritään huomioimaan valtionosuusjärjestelmässä erilaisten valtionosuuskriteereiden avulla.

Merkittävin osa kuntien peruspalvelujen valtionosuudesta muodostuu laskennallisesta valtionosuudesta, joka perustuu kunnan väestön ikärakenteen mukaisiin laskennallisiin kustannuksiin ja muihin laskennallisiin kustannuksiin. Ikärakennetekijän paino laskennallisista kustannuksista on yli 80 prosenttia. Loput laskennallisista kustannuksista määräytyvät kunnan työttömyysasteen, vieraskielisyyden, kaksikielisyyden, saaristoisuuden, asukastiheyden, ja koulutustaustan perusteella. Lisäksi valtionosuuskriteereihin kuuluu lisäosia, joilla on kustannustasauksen ohella tavoiteltu myös ohjausvaikutusta tai muita tavoitteita.

Kunnan peruspalvelujen valtionosuudesta annetun lain¹ (618/2021) 6–17 §:ssä tarkoitetut perushinnat vuonna 2024 on kuvattu Taulukossa 1.

Taulukko 1. Valtionosuuskriteerit ja niiden perushinnat vuonna 2024

Ikäluokittaiset perushinnat	euroa/asukas
0–5-vuotiaat	8 186,31
6-vuotiaat	8 686,50
7–12-vuotiaat	7 231,33
13–15-vuotiaat	12 434,58
16 vuotta täyttäneet ja tätä vanhemmat	64,06
Muut perushinnat	euroa/asukas
Työttömyyskerroin	69,27
Vieraskielisyys	1 717,54
Kaksikielisyyys	294,01
Saaristoisuus; saaristokunta	404,68
Saaristoisuus; saaristo-osakunta	296,02
Asukastiheyserroin	41,53
Koulutustaustakerroin	28,43

¹ <https://www.finlex.fi/fi/laki/ajantasa/2021/20210618#P6>

Syrjäisyyden lisäosa	62,59
Työpaikkaomavaraisuuden lisäosa	13,10
Hyvinvoinnin ja terveyden edistämisen lisäosa	19,31
Asukasmäärän kasvun lisäosa	10,24
Saamelaisten kotiseutualueen kunnan lisäosa	915,13

Selite: Taulukossa näytetään nykyisen valtiosuusjärjestelmän valtiosuuskaiteerit ja niiden perushinnat vuonna 2024.

Perushintojen määritelmät ja käyttö valtiosuuden laskennassa kuvataan sanallisesti laissa². Useimpien valtiosuuskaiteerien vaikutus kunnan asukaskohtaisen valtiosuuteen lasketaan suoraviivaisesti kaiteerin arvon ja perushinnan tulona, mutta eräiden kaiteerien soveltamiseen liittyy lisäehtoja. Selvyyden vuoksi alla kuvataan kunkin valtiosuuskaiteerin vaikutusta kunnan valtiosuuteen asukasta kohden yhtälöiden avulla.

Ikäluokat

Euromäärä/asukas = perushinta*ikäryhmän väestöosuus

Työttömyyskerroin

Kunnan työttömyyskerroin lasketaan jakamalla kunnan työttömien osuus kunnan työvoimasta koko maan vastaavalla osuudella. Kerrointa laskettaessa käytetään työ- ja elinkeinoministeriön työnvälitystilaston tietoa varainhoitovuotta edeltävää vuotta edeltäneen vuoden työttömien osuudesta työvoimasta.

Euromäärä/asukas = perushinta*työttömyyskerroin

Vieraskielisyys

Euromäärä/asukas = perushinta*vieraskielisten osuus

Kaksikielisyys

Euromäärä/asukas= 0.07*perushinta + 0,93*ruotsinkielisten osuus * perushinta
jos kunta on kaksikielinen

Saaristoisuus (saaristokunta)

Euromäärä/asukas

² <https://www.finlex.fi/fi/laki/ajantasa/2021/20210618#L2P10>

= perushinta	jos saaristokunta ja alle 50% väestöstä asuu ilman kiinteää tieyhteyttä mantereeseen
= 3* perushinta	jos saaristokunta ja yli 50% väestöstä asuu ilman kiinteää tieyhteyttä mantereeseen

Saaristoisuus (saaristo-osakunta)

Euromäärä/asukas

= perushinta*saaristossa asuvien osuus	jos saaristo-osakunta
--	-----------------------

Asukastiheyskerroin

Kunnan asukastiheyskerroin lasketaan jakamalla koko maan keskimääräinen asukastiheys kunnan asukastiheydellä. Kerroin on kuitenkin enintään 20.

Euromäärä/asukas = perushinta*asukastiheyskerroin

Koulutustaustakerroin

Kunnan koulutustaustakerroin saadaan laskemalla ilman perusasteen jälkeistä tutkintoa olevien 30–54-vuotiaiden asukkaiden osuus kunnan vastaavasta ikäryhmästä ja jakamalla se koko maan vastaavalla osuudella.

Euromäärä/asukas = perushinta*koulutustaustakerroin

Syrjäisyys

Euromäärä/asukas = perushinta*syrjäisyysluku	jos syrjäisyysluku<1
= 1,5*perushinta* syrjäisyysluku	jos 1<syrjäisyysluku<1,5
= 3*perushinta* syrjäisyysluku	jos 1,5<syrjäisyysluku

Työpaikkaomavaraisuus

Työpaikkaomavaraisuuskerroin saadaan jakamalla kunnan työpaikkojen osuus kunnan työllisten määrästä koko maan vastaavalla osuudella.

Euromäärä/asukas = perushinta*työpaikkaomavaraisuuskerroin

Hyvinvoinnin ja terveyden edistäminen

Kunnan hyvinvoinnin ja terveyden edistämisen kerroin määritellään kunnan hyvinvoinnin ja terveyden edistämisen toimintaa kuvaavien indikaattoreiden sekä toiminnan tulosta kuvaavien indikaattoreiden perusteella. Tähän lisäosaan viitataan tässä tutkimuksessa hyte-lisäosana.

Euromäärä/asukas = perushinta*hyte-kerroin

Asukasmäärän kasvu

Asukasmäärän kasvun kerroin on kolmen vuoden asukasmäärän vuosittaisten muutosprosenttien keskiarvo.

Euromäärä/asukas = perushinta*asukasmäärän kasvun kerroin jos kerroin>0

Saamelaisuus

Euromäärä/asukas = perushinta*saamenkielisten osuus jos saamelaisten kotiseutualue

2.2. Nykyisten valtionosuuskriteerien huomioiminen tässä tutkimuksessa

Nykyisen valtionosuusjärjestelmän perusrakenne otetaan tämän tutkimuksen empiirisen analyysin lähtökohdaksi. Empiirisessä analyysissä käytetään lineaarista funktiomuotoa, joka mahdollistaa regressiomallien kertoimien hyödyntämisen suoraan valtionosuuskriteerien euromääräisinä perushintoina.

Myös tilastomalleissa käytettävien kustannuksia selittävien muuttujien valinnassa lähtökohtana on nykyisen valtionosuusjärjestelmän kriteerit. Kaikkia nykyisen järjestelmän kriteerejä ei kuitenkaan käytetä tarkastelussa sellaisessa muodossa kuin kriteerit on sisällytetty nykyiseen järjestelmään. Lisäksi nykyisen järjestelmän kriteereistä poistetaan kriteerit, joiden sisällyttämiselle valtionosuusjärjestelmään ei ole selkeää teoreettista perustelua. Seuraavaksi esittelemme ja perustelemme nykyisten kriteerien ja tämän tutkimuksen tilastoanalyysissä käytettyjen muuttujien eroja.

Nykyisen valtionosuusjärjestelmän kunnan vieraskielisten määrää kuvaava muuttuja on korvattu syntyperä ulkomailla -muuttujalla. Syntyperältään ulkomaalaisiksi katsotaan henkilöt, joiden molemmat vanhemmat tai ainoa vanhempi on syntynyt ulkomailla. Syntyperän hyödyntämistä voidaan pitää vieraskielisyyttä parempana tapana mallintaa puutteellisesta kotimaisten kielten taidosta aiheutuvia lisäkustannuksia, sillä vieraskielisyys perustuu todellisen kielitaidon sijasta omaan ilmoitukseen äidinkielestä. Moni äidinkieleltään vieraskielinen saattaa puhua sujuvaa suomea ja on voinut suorittaa peruskoulun Suomessa.

Nykyisessä valtionosuusjärjestelmässä huomioidaan kuntien mahdollinen kaksikielisyys siten, että kaksikielisten kuntien saama lisäys valtionosuuteen kasvaa ruotsinkielisten osuuden kasvaessa. Korvaamme tarkasteluissa ruotsinkielisten osuuden vähemmistökielisten osuudella kaksikielisessä kunnassa. Vähemmistökielisten väestöosuus on paremmin perusteltu, sillä ruotsinkielisten osuuden kasvaessa yli 50 prosentin, ei osuuden kasvu enää oletettavasti aiheuta lisäkustannuksia.

Saaristoisuus huomioidaan nykyisissä valtionosuuskriteereissä saaristokuntia kuvaavan indikaattorin avulla ja perushinta lasketaan kolminkertaisena, kun yli puolet väestöstä on vailla kiinteää tieyhteyttä mantereeseen. Lisäksi nykymallissa valtionosuuteen vaikuttaa saaristossa asuvan väestön osuus saaristo-osakunnissa. Nykyiset kriteerit voivat olla osittain päällekkäisiä syrjäisyysmuuttujan kanssa. Saaristokuntien valtionosuuden epäjatkuva funktiomuoto ei myöskään vaikuta perustellulta. Nykyiset kriteerit korvataan analyysissä saaristoliikenteen varassa tai ilman kiinteää yhteyttä olevan väestön väestöosuudella, joka voidaan nähdä mittaavan saaristoisuudesta

aiheutuvia hankalien liikenneyhteyksien lisäkustannuksia, joita syrjäisyysluku ei kuvaa, ilman että kuntia jaetaan epäjatkuviin luokkiin.

Työttömyysastetta, koulutustaustakerrointa, työpaikkaomavaraisuutta ja väestönkasvuastetta ei sisällytetä päätuloksiksi luokiteltaviin tarkasteluihin, sillä niiden sisällyttämisellä ei ole selkeää teoreettista perustaa sote-uudistuksen jälkeisessä tilanteessa, jossa valtionosuuksien kustannuspohja koostuu varhaiskasvatuksen, koulutuksen ja kulttuuripalvelujen kustannuksista. Päätulosten herkkyyksianalyysissa edellä mainitut tekijät lisätään tarkasteluun työpaikkaomavaraisuutta lukuun ottamatta. Tarkastelujen läpinäkyvyyden vuoksi on syytä testata, kuinka vahva yhteys näillä nykyjärjestelmän vos-kriteereillä on kuntien menojen kanssa. Vuonna 2025 voimaan tulevan TE-uudistuksen myötä kuntien työmarkkinoiden tilan ja kuntien TE-menojen yhteyden tarkastelu on jatkossa tarpeellista.

Hyvinvoinnin ja terveyden edistämisen mittari on rajattu tarkastelun ulkopuolelle, sillä hyte-lisäosa perustuu yhdistelmään hyvinvoinnin ja terveyden edistämisen toimintaa kuvaavia sekä toiminnan tulosta kuvaavien indikaattoreita.

3. Aineisto ja menetelmät

3.1. Tutkimusaineisto

Tutkimuksessa käytetään kuntatason aineistoja palvelutuotannon toimintamenoista ja väestön ikärakenteesta sekä kunnan muista ominaisuuksista vuosilta 2018–2022.

Palveluluokittaiset kustannustiedot on poimittu vuosilta 2018–2020 Tilastokeskuksen kuntataloustilastoista ja vuosilta 2021–2022 Valtiokonttorin tutkiahallintoa.fi palvelusta. Näistä tilastoista tutkimukseen sisällytetään toimintakulut palveluluokista, jotka vastaavat mahdollisimman hyvin tarkasteltavia palveluita (varhaiskasvatus, esi- ja perusopetus ja kirjastot, taiteen perusopetus ja yleinen kulttuuritoimi). Kunnan väestöä, sen ikärakennetta ja muita kunnan ominaisuuksia koskevat tiedot poimitaan Tilastokeskuksen muista tilastoista sekä muista lähteistä.

Tutkimusaineiston toimintamenoihin tehdään muutoksia, joiden tavoitteena on parantaa kuntien välistä vertailukelpoisuutta. Nämä muokkaukset kuvataan liitteessä A. Muutosten takia toimintamenot eivät täysin vastaa tilastolähteistä saatavia tietoja. Taulukko 2 havainnollistaa vuosittaisten kokonaistoimintamenojen arvoa muutoksien kanssa ja ilman muutoksia vuosina 2018–2022.

Taulukko 2. Kuntien toimintamenojen kustannuspohja vuosina 2018–2022 (reaalinen 2022 rahassa)

	Toimintamenot muutoksien kanssa yhteensä, miljoonaa euroa	Toimintamenot yhteensä ilman muutoksia, miljoonaa euroa
2018	9959,21	10013,64
2019	10208,19	10259,9
2020	10098,38	10135,3
2021	10230,83	10255,28
2022	10235,95	10271,66

Selite: Taulukossa on tutkimuksessa tarkasteltavien toimintakulujen summa vuosina 2018–2022 vertailukelpoisuutta parantavien muokkausten jälkeen (vasen sarake) ja ennen muokkauksia (oikea sarake).

Taulukon 2 toisessa sarakkeessa esitellään kustannuspohja, johon tämän raportin tarkastelut perustuvat. Kustannuspohjalla tarkoitetaan kuntien yhteenlaskettuja toimintamenoja. Taulukon 2 kolmannessa sarakkeessa raportoidaan vertailun vuoksi kustannuspohja, josta ei ole poistettu tai lisätty kustannuksia. Vertailtavuuden parantamiseksi vuoksi sarakkeiden 2 ja 3 kustannukset on deflatoitu samalla hintaindeksillä vuoden 2022 arvoon. Tutkimuksessa tarkasteltavat toimintamenot ovat noin kymmenen miljardia euroa vuodessa. Muutoksien seurauksena toimintamenot ovat noin 30–50 miljoonaa euroa pienemmät kuin toimintamenot ilman muutoksia. Toimintamenoihin tehtävien muutoksien mittakaava on kustannuspohjaan verrattuna hyvin pientä, koska osa muutoksista pienentää kustannuspohjaa ja osa kasvattaa sitä. Taulukosta nähdään lisäksi, että kustannuspohjassa ei tapahdu merkittävää muutosta, kun toimintamenojen aineistolähde muuttuu Tilastokeskuksesta Valtionkonttoriksi vuonna 2021.

Taulukon 2 kustannuspohja poikkeaa jonkin verran peruspalvelujen valtionosuuksien kustannuspohjasta. Esimerkiksi vuonna 2023 kustannuspohja oli VM:n valtionosuuslaskelman mukaan noin 9,6 miljardia euroa.

Taulukossa 3 esitellään tutkimusaineiston tilastolliset tunnusluvut regressiomalleihin sisällytettyjen muuttujien osalta vuosina 2018–2022. Taulukon tunnusluvut on laskettu otoksesta, jota myöhemmin käytetään tilastollisessa regressioanalyysissä. Havaintoyksikkö on kunta tietynä vuonna. Muuttujien muodostaminen, muokkaaminen ja aineistolähteet kuvataan liitteessä A. Aineistosta on pudotettu Ahvenanmaan kunnat sekä kunnat, joilta jonkin muuttujan arvo puuttuu. Lisäksi Savonlinnan vuosien 2021 ja 2022 havainnot pudotetaan, koska kunnan toimintamenot näiltä vuosilta sisältää mittavirhettä. Aineiston havaintomäärä on 292 kuntaa vuonna 2018, 293 kuntaa vuosina 2019–2020 ja 291 kuntaa vuosina 2021–2022.

Taulukon 3 ensimmäisellä rivillä on väestöön suhteutettuja toimintamenoja kuvaava muuttuja ja muilla riveillä on regressioanalyysissä selittävinä muuttujina käytettävät muuttujat. Toimintamenot on muunnettu reaalisiksi vuoden 2022 tasoon kuntien peruspalvelujen hintaindeksiä käyttämällä. Taulukon sarakkeissa on kunkin muuttujan keskiarvo, keskihajonta, pienin arvo ja suurin arvo.

Taulukosta 3 nähdään, että asukaskohtaisissa toimintamenoissa on merkittävää kuntien välistä vaihtelua. Vuosien 2018–2022 välillä keskimääräinen kustannus oli noin 1900 €/asukas, mutta

samalla pienin ja suurin kustannus olivat 1130 €/asukas ja 3733€ €/asukas. Myös ikäryhmien väestöosuuksissa on suuria eroja. Esimerkiksi 0–5-vuotiaiden väestöosuus vaihtelee 1,2 prosentin ja 12,3 prosentin välillä. Myös esimerkiksi syrjäisyyttä ja saaristoisuutta mittaavissa muuttujissa keskihajonta on muuttujan keskiarvoa suurempaa, mikä kertoo kyseisten muuttujien arvojen vaihtelevan merkittävästi. Kunnan sosioekonomisten ominaisuuksien (työttömyysaste ja ei tutkintoa osuus) tai kunnan väestön syntyperän osalta keskihajontana mitattu vaihtelu kuntien välillä ovat selkeästi pienempää.

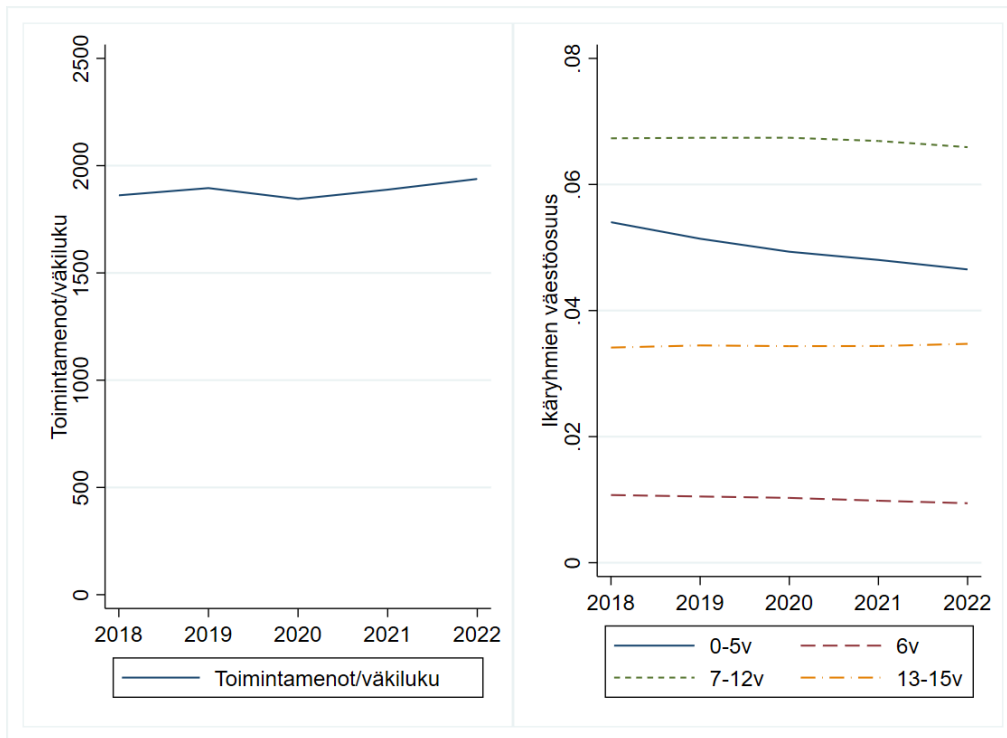
Kuvio 1 havainnollistaa reaalisten toimintamenojen ja ikäryhmien väestöosuuksien keskiarvojen kehitystä vuosina 2018–2022. Kuvion vasemmassa paneelissa havaitaan keskimääräisten menojen kehittyvän varsin vakaasti. Kustannustiedoissa aineistolähde vaihtuu vuonna 2021 Tilastokeskuksesta Valtiokonttoriin, mutta keskiarvon kehityksessä ei tapahdu tuolloin merkittävää muutosta. Kuvion 1 oikea paneeli näyttää ikäryhmien keskimääräisten väestöosuuksien kehityksen. 0–5-vuotiaiden osuuden keskiarvo laskee ajanjaksolla noin prosenttiyksiköllä. Muiden ikäryhmien keskiarvot pysyvät lähes ennallaan tarkastelujakson ajan.

Taulukko 3. Regressiomallien muuttujien tilastolliset tunnusluvut vuosilta 2018-2022

	Keskiarvo	Keski-hajonta	Min	Max
Toimintamenot / väestöpohja	1886	334	1130	3733
0-5-vuotiaiden väestöosuus	0,050	0,015	0,012	0,123
6-vuotiaiden väestöosuus	0,010	0,003	0,001	0,026
7-12-vuotiaiden väestöosuus	0,067	0,018	0,026	0,147
13-15-vuotiaiden väestöosuus	0,034	0,008	0,013	0,072
Kaksikielinen kunta	0,113	0,317	0,000	1,000
Vähemmistökielisten osuus	0,026	0,086	0,000	0,479
Syntyperä ulkomailla, väestöosuus	0,035	0,029	0,004	0,253
Saame kotiseutu kunta	0,014	0,116	0,000	1,000
Syrjäisyysluku, tie	0,531	0,624	0,000	1,954
Saaristo-osuus	0,006	0,059	0,000	0,984
Työttömyysaste	0,107	0,034	0,020	0,222
Ei tutkintoa osuus	0,140	0,035	0,050	0,260
Väestönkasvu	-0,008	0,012	-0,066	0,046
Laskennallinen veropohja/väestö	3336	618	2035	8392
Kunnan väestö (1000 henkeä)	18,8	51	0,7	664
Kunnan väestö² (1000 henkeä)	2953	26274	0,5	440933

Selite: Taulukossa näytetään regressioanalyysissä käytettävän aineiston muuttujien tilastolliset tunnusluvut vuosilta 2018–2022. Havaintojen määrä on 1460.

Kuvio 1. Asukaskohtaisten toimintakulujen ja ikäryhmien väestöosuksien kehitys vuosina 2018–2022.



Selite: Kuviossa näytetään kuntien väkiluvulla jaettujen toimintakulujen keskiarvon (vasen kuva) ja valtiosuusjärjestelmässä käytettyjen ikäryhmien väestöosuuksien keskiarvojen (oikea kuva) kehitys vuosina 2018–2022. Havaintojen määrä on 1460.

3.2. Tutkimusmenetelmä

Tutkimusmenetelmänä käytetään tilastollista regressioanalyysia, jossa selitettävä muuttuja on tarkasteltavien palveluiden toimintakulut asukasta kohti. Selittävinä muuttujina käytetään eri ikäryhmien väestöosuuksia ja muita kunnan ominaisuuksia kuvaavia muuttujia. Analyysissa käytetyt ikäryhmät vastaavat nykyisen valtiosuusjärjestelmän ikäryhmiä. Muita kunnan ominaisuuksia kuvaavat muuttujat voivat sisältää esimerkiksi nykyisiä valtiosuuskkriteereitä tai uusia valtiosuuskkriteereiksi harkittavia muuttujia. Mallispesifikaatioiden valinnassa ja tulosten tulkinnassa hyödynnetään aiempaa VATT:n tutkimusta kuntien valtiosuuskkriteereistä (ks. Harjunen ja Saarimaa, 2012).

Regressioanalyysissa aineistona käytetään vuodet 2018–2022 kattavaa paneeliaineistoa, mutta kuntakohtaisia kiinteitä vaikutuksia ei kontrolloida sillä osa kuntien ominaisuuksista vaihtelee kunnan sisällä ajan myötä vain vähän tai ei lainkaan (esim. kaksikielisyys tai saamelaisuus), jolloin muuttujien kertoimia ei voida estimoida kiinteiden vaikutusten mallilla lainkaan tai estimaatit ovat liian epätarkkoja. Regressiomallien lisäksi tilastollista tarkastelua täydennetään laskemalla havaintoaineistosta suoraan keskimääräisiä kustannuksia erilaisille ikäryhmä- ja kustannusluokille.

Tutkimuksessa käytetyt pienimmän neliösumman regressiomallit ovat muodoltaan seuraavanlaisia:

$$(1) \quad \frac{C_{it}}{N_{it}} = \alpha_t + \sum_{k \neq 5}^K \beta_k \frac{n_{itk}}{N_{it}} + \delta' z_{it} + \epsilon_{it} .$$

Mallin selitettävä muuttuja $\frac{C_{it}}{N_{it}}$ on väestöön suhteutetut toimintamenot kunnassa i kalenterivuonna t . Kunnan asukasluku N_{it} vastaa kaikkien ikäryhmätermien n_{itk} summaa $\sum_{k=1}^K n_{itk}$. Muuttujat $\sum_{k \neq 5}^K \beta_k \frac{n_{itk}}{N_{it}}$ kuvaavat malliin sisällytettyjä väestöryhmäosuuksia kertoimineen. Mallin ulkopuolelle jätettävä ikäryhmä $k = 5$ on yli 15-vuotiaiden ikäryhmä. Muut ikäryhmät ovat nuorimmasta vanhimpaan ikäryhmät 1-4. Mallin vakiotermin α_t sallitaan vaihdella vuosien välillä sisällyttämällä malliin vuosikohtaiset kiinteät vaikutukset, jotka kontrolloivat kaikille kunnille yhteiset ajassa tapahtuvat kustannusmuutokset. Termi z_{it} sisältää muut mahdolliset valtionosuuskriteerit. Lisäksi z_{it} sisältää kontrollimuuttujina käytettäviä kunnan ominaisuuksia. Mallin virhetermi on ϵ_{it} .

Kontrollimuuttujina käytetään kunnan laskennallista veropohjaa asukasta kohden, väkilukua ja väkiluvun toisen asteen termiä sekä kaksivuotisen esikoulun kokeiluun osallistuneiden lasten väestöosuutta vuonna 2022. Veropohjan avulla pyritään vakioimaan kuntalaisten tulotason ja vaurauden vaikutus palvelutasoon. Väkiluvun kontrolloinnilla puolestaan pyritään ottamaan huomioon skaalaedut palvelutuotannossa. Esikoulukokeiluun osallistuneiden osuus kontrolloidaan sillä esikoulukokeilusta vuonna 2022 aiheutui lisämenoja, jotka eivät kuulu normaaliin peruspalvelujen kustannuspohjaan.

Luvussa 4 estimoidaan kolme erilaista regressiomallia, jotka eroavat toisistaan termiin z_{it} sisällytettävien muuttujien suhteen. Yksinkertaisimpaan malliin sisällytetään ikäryhmien lisäksi vain kontrollimuuttujat ja kalenterivuotta kuvaavat kiinteät vaikutukset. Toiseen malliin lisätään kuntien kieliasemaa kuvaavia muuttujia. Kolmannessa regressiomallissa selittäviksi muuttujiksi tuodaan kunnan syrjäisyyttä kuvaavia muuttujia. Liitteessä A kuvataan muuttujat ja muuttujien aineistolähteet tarkemmin.

Regressiomallien keskivirheet ryvästetään (klusteroidaan) kuntatasolla. Tällöin keskivirheiden laskennassa sallitaan kunnan sisäisiä esimerkiksi ajallisia riippuvuuksia selittämättä jäävissä toimintamenoissa (virhetermi), mutta kuntien virhetermien oletetaan olevan riippumattomia muiden kuntien virhetermeistä.

Regressioanalyysin kerroinestimaatit kuvaavat selittävien muuttujien keskimääräistä yhteyttä tutkittavien palveluiden kustannuksiin. Tuloksia voidaan hyödyntää valtionosuusjärjestelmän kriteerien valinnassa ja niiden perushintojen määrittämisessä. Tässä tutkimuksessa estimoituja kertoimia voidaan pääosin pitää nykyjärjestelmän perushintoja parempina arvioina siitä, miten vos-kriteerit vaikuttavat kunnan toimintamenoihin. Nykyiset vos-kriteerit perustuvat vanhentuneisiin selvityksiin kuntien kustannusten ja vos-kriteerien yhteydestä, sekä asiantuntijoiden ja poliitikkojen näkemyksiin. Tämän tutkimuksen tulokset tarjoavat objektiivisen ja aiempia selvityksiä ajantasaisemman kuvan sote-uudistuksen jälkeisten palvelujen menoja selittävistä tekijöistä.

Regressiokertoimien ei kuitenkaan voida tulkita kuvaavan kovinkaan suurella varmuudella selittävien muuttujien kausaalista vaikutusta kuntien menoihin. Syyseuraussuhteiden luotettavampaan arviointiin tarvittaisiin ns. luonnollisia koeasetelmia, jotka aiheuttaisivat eksogeenista variaatiota tutkittavissa vos-kriteereissä. Tällaista tutkimusasetelmaa ei ole tarjolla, jolloin tämän tutkimuksen malli on hyvin perusteltu tapa pyrkiä arvioimaan vos-kriteerien vaikutuksia kustannuksiin. Muuttujien kertoimet voivat olla harhaisia arvioita muuttujien vaikutuksista menoihin, jos muuttujat korreloivat

mallin ulkopuolisten tekijöiden kanssa, jotka sisältyvät virhetermiin. Lisäksi muuttujiin saattaa sisältyä mittausvirheitä, joka myös aiheuttaa harhaa tuloksiin. Esimerkiksi ikäryhmät on muodostettu kunnan ikärakenteen eikä todellisen palvelukäytön perusteella.

Regressioyhtälön (1) estimoidut kertoimet (α , β ja δ) kertovat, kuinka yhden yksikön lisäykset mallin selittävässä muuttujissa heijastuvat kunnan väestöön suhteutettuihin toimintamenoihin. Näiden kertoimien avulla voidaan laskea, kuinka paljon yhden henkilön lisääminen mallin väestöryhmiin lisää kunnan toimintamenoja. Ikäryhmien laskennalliset yksikkökustannukset voidaan laskea seuraavasti:

$$(2) \quad \begin{cases} \text{Ryhmän 5 yksikkökustannus : } \widehat{\alpha}_{2022} + \delta' \bar{z} \\ \text{Ryhmän } k \text{ yksikkökustannus : } \widehat{\alpha}_{2022} + \delta' \bar{z} + \widehat{\beta}_k, \quad k \neq 5 \end{cases}$$

Regressiomallissa vertailuryhmänä käytetyn ryhmän 5 (yli 15-vuotiaiden ryhmä) yksikkökustannus perustuu regressiomallin vakiotermiin $\widehat{\alpha}_{2022}$ ja kontrollimuuttujien $\delta' \bar{z}$ kertoimiin, ja muiden ryhmien yksikkökustannuksiin vaikuttaa myös estimoitu ikäryhmämuuttujan kerroin $\widehat{\beta}_k$. Yksikkökustannuksia laskettaessa mallissa olevien valtiosuusjärjestelmään sisällytettävien muuttujien arvot asetetaan nolnaan, jottei niitä huomioida moneen kertaan. Kontrollimuuttujien arvot puolestaan asetetaan keskiarvoon. Yhtälön (2) rakennetta käyttäen voidaan laskea ikäryhmittäiset yksikkökustannukset erilaisissa regressiomalleissa. Yksikkökustannuksille lasketaan myös niiden tilastollista epävarmuutta kuvaava keskivirhe.

Päätuloksissa on hyödynnetty vuosien 2018–2022 aineistoa, mutta herkkyytarkasteluissa tarkastellaan aikajakson muutosten vaikutuksia tuloksiin. Aikajakson valinta on tasapainoilua estimoitujen parametrien tarkkuuden ja aineiston ajantasaisuuden välillä. Tarkasteluvuosien vähentäminen voi kasvattaa estimoitujen parametrien keskivirheitä, mutta toisaalta esimerkiksi vuoden 2018 aineistossa kustannusten ja selittävien muuttujien yhteys ei välttämättä vastaa hyvin nykyhetken tilannetta. Aineiston rajaaminen esimerkiksi vuosiin 2021–2022 mahdollistaa sen, että aineisto on hyvin ajantasaista, mutta tällöin tilastollinen epävarmuus kasvaa. Muuttujien kertoimien kehitys ajanjaksoa muutettaessa antaa viitteitä siitä, onko kustannusten ja selittävien muuttujien yhteys pysynyt ajan myötä vakaana. Herkkyytarkasteluissa testataan lisäksi useita vaihtoehtoisia malleja, joissa päämallin muuttujiin on tehty muutoksia.

4. Kuntien toimintamenojen tarkastelu

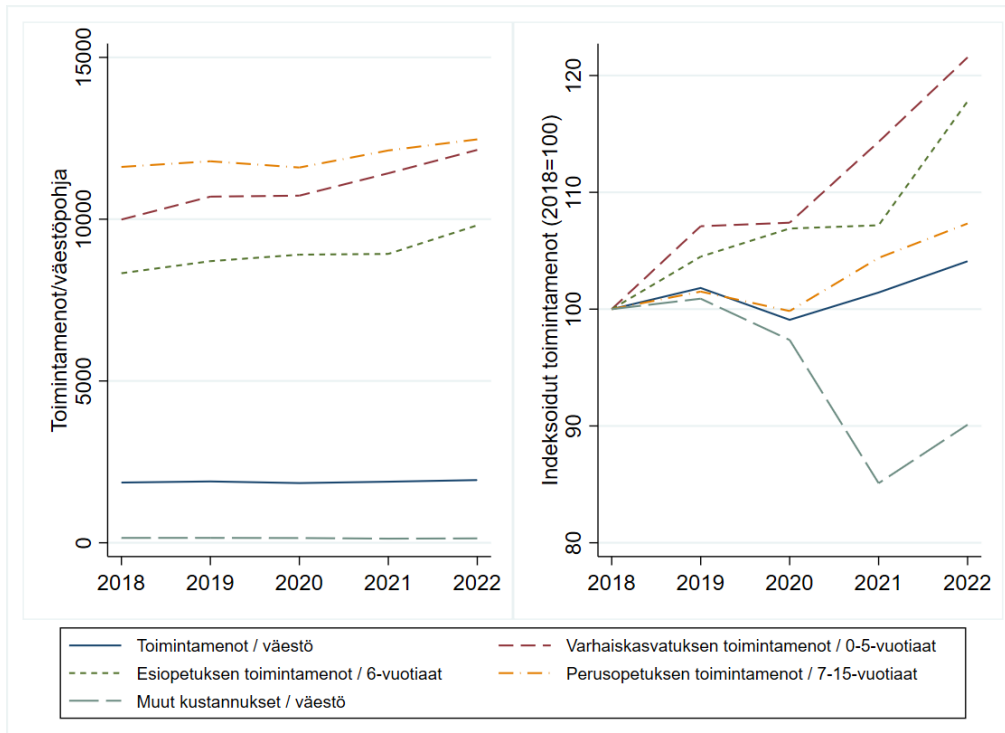
Tässä osiossa tarkastellaan aluksi kuntien keskimääräisten väestöpohjaan suhteutettujen toimintamenojen kehitystä kokonaisuutena sekä palveluluokittain. Seuraavaksi tarkastellaan regressioanalyysin avulla, miten kuntien väliset erot mahdollisissa valtiosuusskiteereissä heijastuvat toimintamenoihin.

4.1. Kuntien toimintamenojen kuvaileva tarkastelu

Kuvio 2 havainnollistaa kuntien toimintamenojen kehitystä vuosina 2018–2022. Asukaskohtaisten kokonaistoimintamenojen lisäksi tarkastellaan myös palveluluokittaisia toimintamenoja palvelun piirissä olevaan ikäluokkaan kuuluvaa kohden. Omiksi kustannusluokiksi on eroteltu varhaiskasvatus, esiopetus, perusopetus ja muut kustannukset. Muut kustannukset palveluluokka

sisältää kaikki edellä mainittujen kustannusluokkien ulkopuoliset kustannukset. Kuvan vasen paneeli näyttää toimintamenojen kehityksen euromääräisenä (reaalinen). Oikea paneeli näyttää vastaavien toimintamenojen kehityksen indeksoituna vuoteen 2018.

Kuvio 2. Kuntien toimintamenojen kehitys vuosina 2018–2022



Selite: Kuviossa näytetään eri palveluluokkien reaalisten väestöpohjaan suhteutettujen toimintamenojen keskiarvon kehitys vuosina 2018–2022 euromääräisenä (vasen kuvio) ja indeksoituna (oikea kuvio).

Kuviosta 2 havaitaan kokonaismenojen pysyneen varsin vakaina vuodesta 2018 vuoteen 2022. Kehityksessä on kuitenkin palveluluokkien välisiä eroja. Palveluluokka muut kustannukset on ainoa, jossa tapahtuu laskua verrattuna vuoteen 2018. Muiden palveluluokkien menoilla on kasvava trendi. Erityisesti varhaiskasvatuksen ja esiopetuksen toimintamenot ovat kasvaneet vuodesta 2018.

Taulukossa 4 raportoidaan tarkemmin Kuvan 2 vasemman paneelin vuosittaiset arvot. Taulukon ensimmäisellä rivillä on kuntien väestöön suhteutetut kokonaistoimintamenojen kehitys vuosina 2018–2022. Seuraavilla kolmella rivillä esitellään vastaavasti varhaiskasvatuksen, esiopetuksen ja perusopetuksen palveluiden piirissä olevaan väestöön suhteutetut menot. Viimeisellä rivillä esitellään varhaiskasvatuksen, esiopetuksen ja perusopetuksen ulkopuoliset kunnan väestöön suhteutetut muut menot. Reaaliset kokonaistoimintamenot asukasta kohden ovat nousseet vain hieman vuosien 2018–2022 välillä. Nousua on tapahtunut varhaiskasvatuksen, esiopetuksen ja perusopetuksen menoissa, kun taas muut menot ovat laskeneet. Muut menot ovat väestöpohjaa kohti laskettuna tasoltaan pieni erä, mutta väestöpohjaan suuruuden takia silläkin on suuri merkitys kokonaismenojen kannalta.

Taulukko 4. Kuntien keskimääräiset vuosittaiset toimintamenot palveluluokittain.

Toimintamenuoluokka	2018	2019	2020	2021	2022
<i>Toimintamenot</i>	1862,7	1895,5	1845,0	1890,3	1940,1
<i>Kunnan kokonaisväestö</i>	(315,5)	(319,9)	(309,4)	(348,3)	(361,3)
<i>Varhaiskasvatuksen toimintamenot</i>	9988,2	10697,4	10727,5	11418,9	12140,8
<i>0 – 5 vuotiaat</i>	(2280,2)	(2307,4)	(2018,9)	(2237,9)	(2216,5)
<i>Esiopetuksen toimintamenot</i>	8329,1	8703,5	8903,9	8922,7	9799,6
<i>6 vuotiaat</i>	(3626,8)	(4700,2)	(1098,0)	(4334,9)	(4101,0)
<i>Perusopetuksen toimintamenot</i>	11615,2	11793,8	11596,5	12129,0	12470,6
<i>7 – 15 vuotiaat</i>	(2678,3)	(2661,5)	(2547,5)	(3117,1)	(3196,2)
<i>Muut toimintamenot</i>	150,5	151,8	146,5	128,1	135,6
<i>Kunnan kokonaisväestö</i>	(77,1)	(77,4)	(76,8)	(81,6)	(107,9)

Selite: Taulukossa on eri palveluluokkien reaalisten väestöpohjaan suhteutettujen toimintamenojen keskiarvo ja keskihajonta (suluissa) vuosina 2018-2022.

4.2. Kuntien toimintamenojen regressioanalyysi

Taulukossa 5 esitellään regressiomallien tulokset. Kaikissa mallispesifikaatioissa selitetään kuntien asukaskohtaisia toimintakuluja kuntien ikärakenteella ja kuntien muilla ominaisuuksilla. Taulukko sisältää kolmen regressiomallin tulokset, joista ensimmäinen malli on suppein ja kolmas malli sisältää suurimman määrän selittäviä muuttujia. Selittävät muuttujat esitellään raportin luvussa 3.2 ja tarkemmin liitteessä A.

Taulukko 5. Regressioanalyysin tulokset.

Selitettävä muuttuja: Toimintakulut asukasta kohden			
Malli:	M1	M2	M3
0-5 -vuotiaiden osuus	8498.5*** [2120.6]	6787.1*** [1870.9]	6731.5*** [1723.4]
6-vuotiaiden osuus	1401.3 [3415.1]	4301 [2848.6]	6503.8** [2769.3]
7-12-vuotiaiden osuus	3099.5 [2673.2]	3731.4* [1918.2]	5466.1*** [1753.1]
13-15-vuotiaiden osuus	6188.2 [3764.7]	8025.6*** [2834.0]	9305.9*** [2696.8]
Kaksikielinen kunta		63.16 [64.89]	36.27 [62.47]
Vähemmistökielisten osuus		381.8* [230.6]	418.2* [232.0]
Syntyperä ulkomailla osuus		801.7 [505.7]	1202.5** [491.0]
Saame kotiseutu kunta		944.7*** [237.0]	783.7*** [225.5]
Syrjäisyysluku			156.8*** [22.79]
Saaristoliikenne tai ei yhteyttä osuus			249.0** [101.3]
Laskennallinen verotulo/asukas	0.0411* [0.0215]	-0.0145 [0.0227]	0.0312* [0.0179]
Väkiluku (tuhansia)	-2.089*** [0.553]	-1.859*** [0.537]	-1.157*** [0.443]
Väkiluku(tuhansia)2	0.00282*** [0.000780]	0.00242*** [0.000746]	0.00118* [0.000615]
Esikoulukokeiluun osallistujien osuus	4411.6 [6713.4]	2579.7 [5269.1]	6597.3 [4994.6]
Vakiotermi	917.8*** [85.21]	991.6*** [85.26]	559.4*** [85.43]
N	1460	1460	1460
R ²	0.464	0.594	0.641

Selite: Taulukossa näytetään pns-regressiokertoimet ja hakasuluissa niiden keskivirheet. Tähdet kuvaavat tilastollista merkitsevyyttä (***) $p < 0.01$, ** $p < 0.05$ ja * $p < 0.1$).

Taulukossa 5 esitettyjen kerroinestimaattien etumerkit ovat odotetun suuntaisia. Palvelujen kysyntää tai palvelutuotannon yksikkökustannuksia oletettavasti kasvattavat muuttujat saavat positiivisen kertoimen. Asukaslukua kuvaavan kontrollimuuttujan negatiivinen etumerkki johtuu oletettavasti siitä, että väkiluvun kasvaessa palveluja voidaan tuottaa tehokkaammin. Tuloksien valossa väkiluvun kasvamisesta saatavat kustannussäästöt kuitenkin pienenevät kunnan asukasluvun kasvaessa. Myös kunnan tulopohjaa kuvaava laskennallinen verotulo asukasta kohden on yhteydessä korkeampiin toimintamenoihin.

Tulosten hyödyntämisessä on syytä huomioida, että estimoitujen parametrien suuruus vaihtelee merkittävästi estimoitujen mallien välillä. Esimerkiksi 6-vuotiaiden osuuden kerroin on Mallissa 1 1401.3 ja Mallissa 3 vastaavasti 7387.2. Näin suuret muutokset kertoimissa tarkoittavat sitä, että tulokset ovat herkkiä valitulle mallinnusratkaisuille. Samalla on lisäksi myös huomattava, että estimoitujen parametrien keskivirheet ovat suuret, mikä kertoo merkittävästä tilastollisesta epävarmuudesta. Taulukon 5 malleista vos-järjestelmän uudistamisen pohjaksi soveltuu parhaiten Malli 3, joka sisältää laajimman joukon mahdollisia vos-kriteereitä. Suppeammassa malleissa mallin ulkopuolelle jätettyjen kriteerien vaikutus kanavoituu mallin sisältämien kriteerien kautta, jolloin kertoimet eivät heijasta yhtä luotettavasti muuttujien vaikutuksia kustannuksiin.

Ikäryhmät

Taulukossa 6 esitellään Taulukon 5 regressiomalleista lasketut ikäryhmittäiset perushinnat. Perushinnat on laskettu hyödyntämällä luvussa 3.2 esitetyjä kaavoja. Kolme ensimmäistä saraketta vastaavat Taulukon 5 regressiomallien sarakkeita. Taulukon soluissa esitetään regressiomalliin perustuva perushinta ja hakasuluissa perushinnan keskivirhe. Viimeiseksi sarakkeeksi on sisällytetty myös nykyisen valtiosuusjärjestelmän perushinta vuonna 2023. Vertailun helpottamiseksi neljännessä sarakkeessa esitetään Malliin 3 perustuvat perushinnat, jotka on korotettu vuoden 2023 hintatasoa vastaavaksi kuntien peruspalvelujen hintaindeksiä käyttämällä. Valtiosuustehtävissä tapahtuneista muutoksista aiheutuvia muutoksia perushintoihin ei ole huomioitu sarakkeessa 4.

Taulukko 6. Ikäryhmien perushinnat.

	Malli 1	Malli 2	Malli 3	Malli 3, 2023 taso	Nyky-VOS 2023 taso
0–5-vuotiaat	9601 [2132]	7780 [1891]	7443 [1723]	7776	8011
6-vuotiaat	2504 [3427]	5294 [2864]	7215 [2775]	7538	8500
7–12- vuotiaat	4202 [2668]	4725 [1909]	6177 [1757]	6454	7072
13–15- vuotiaat	7291 [3741]	9019 [2794]	10017 [2657]	10465	12129
16+ vuotiaat	1103 [57]	993 [67]	711 [70]	743	63

Selite: Taulukossa raportoidaan Taulukon 5 regressiomallien tulosten mukaiset ikäryhmien perushinnat, jotka on laskettu kaavalla (2) sekä nykyisen valtiosuusjärjestelmän perushinnat. Perushintojen keskivirheet on hakasuluissa.

Taulukossa 6 raportoidut perushinnat vastaavat 0–5-vuotiaiden, 6-vuotiaiden ja 7–12-vuotiaiden osalta melko läheisesti nykyisen VOS-järjestelmän perushintoja, kun huomioidaan perushinnalle laskettu keskivirhe. Keskivirheiden kokoluokka paljastaa sen, että regressiomallien tulosten tapaan myös perushintoihin liittyy merkittävää tilastollista epävarmuutta. 13–15-vuotiaiden perushinta on noin yhden keskivirheen verran pienempi kuin nykyisen vos-järjestelmän perushinta. Suurin haaste Taulukon 5 perushinnoissa liittyy 16+ vuotiaiden perushintaan, joka on yli kymmenkertainen nykyjärjestelmän hintaan verrattuna. Ei ole syytä ajatella, että 16+ vuotiaiden perushinta vastaisi kyseisen ryhmän todellisia kustannuksia, vaan kyseinen perushinta on seurausta siitä, että käyttämässämme tilastomallissa kyseisen ryhmän perushinnan muodostamiseen käytetyt

parametrit sisältävät harhaa. Myös Taulukossa 4 raportoidut laskennalliset toimintamenot kategorialle muut ovat selvästi Taulukon 5 perushintoja alemmat (135,6 euroa vuonna 2022). Vaikuttaa siltä, että vertailuryhmänä toimivan 16+ vuotiaiden ikäluokan perushinnat yliarvioituvat regressioanalyysissä.

Muut mahdolliset vos-kriteerit

Kaksikielisyyttä kuvaavan indikaattorin kerroin on Mallissa 3 positiivien, mutta ei poikkeaa tilastollisesti merkittävästi nolasta. Vähemmistökielisten osuus kaksikielisissä kunnissa kasvattaa tulosten mukaan kustannuksia. Kertoimen perusteella esimerkiksi 10 prosenttiyksikön nousu vähemmistökielisten osuudessa kasvattaa asukaskohtaisia menoja noin 42 eurolla. Kaksikielisten kuntien vähemmistökielen osuus on keskimäärin 0,26 (maksimi 0,48). Kaksikielisyyttä kuvaavien muuttujien kertoimien perusteella voidaan arvioida, että tällaisella kunnalla on arviolta 144 €/asukas (237 €/asukas) kaksikielisyteen liittyviä lisäkustannuksia

Saamelaisten kotiseutualueen kuntaa kuvaavan indikaattorimuuttujan kerroin on suuri (784 euroa/asukas) ja vahvasti merkitsevä. Nykyjärjestelmän mukaista saamenkielisten osuutta saamelaiskunnissa ei ole sisällytetty malliin. Sen kerroin olisi haastava estimoida mallin puitteissa yhdessä muiden muuttujien kanssa, sillä näitä kuntia on vain neljä.

Syntyperältään ulkomaalaisten osuuden kerroin on noin 1200 ja tilastollisesti merkitsevä. Syntyperä ulkomailla muuttujan keskiarvo on 0,035 (maksimi 0,253). Tällaisella kunnalla arviolta 42 €/asukas (304 €/asukas) syntyperältään ulkomaiseen väestöön liittyviä lisäkustannuksia.

Syrjäisyysluvun kerroin on noin 157 ja sekin on tilastollisesti merkitsevä. Nykyjärjestelmässä syrjäisyysluvun kerroin vaihtuu epäjatkovasti tiettyjen raja-arvojen kohdalla. Vastaavaa funktiomuotoa ei voitu estimoida riittävällä tarkkuudella tässä tutkimuksessa. Syrjäisyysluvun keskiarvo on 0,531 ja maksimi 1,954. Regressiokertoimen perusteella keskiarvon mukaisella kunnalla on syrjäisyyteen liittyviä lisäkustannuksia arviolta 78 €/asukas ja maksimin mukaisella kunnalla 487 €/asukas.

Saaristoliikenteen varassa tai ilman kiinteää liikenneyhteyttä olevan väestön osuuden kerroin on 249.0 (keskivirhe 101.3). Saaristoisuusmuuttujan keskiarvo on 0,006 (maksimi 0,984). Kerroin tarkoittaisi, että tällaisella kunnalla on arviolta 1,5 €/asukas (maksimi 245 €/asukas) saaristoisuuteen liittyviä lisäkustannuksia.

4.3. Regressioanalyysien herkkyystarkastelu

Edellisessä osiossa esiteltiin päämalleiksi valittujen regressiomallien tulokset sekä tuloksista lasketut ikäryhmien perushinnat. Seuraavaksi tarkastellaan, miten tulokset muuttuvat, jos regressiomallien selittäviä tekijöitä muutetaan tai regressiomallien estimointiin käytettyä otosta muutetaan. Herkkyyksanalyysiä suoritetaan siksi, että malliin sisällytetyjä tekijöitä voi mallintaa monella tavalla ja tarkastelun läpinäkyvyyden vuoksi on tärkeää myös havainnollistaa, miten muutokset mallin rakenteessa heijastuvat tuloksiin. Lisäksi osa malliin sisällytetyistä muuttujista on muodostettu niin, että muuttujan arvot vuosille 2018–2020 perustuvat vuoden 2021 arvoon, joten on syytä tarkastella tästä aiheutuvan mittausvirheen mahdollista vaikutusta tuloksiin.

Herkkyystarkasteluja tehdään otokseen sisällytetyjen vuosien suhteen (Taulukko 7), kontrollimuuttujien roolin suhteen, sosioekonomisten tekijöiden lisäämisen suhteen ja asukasluvulla painottamisen suhteen (Taulukko 9), saaristoisuuden määritelmien suhteen (Taulukko 10), saamelaisten kotiseutualuetta kuvaavien muuttujien suhteen (Taulukko 11) ja lopuksi tarkastellaan

asukastiheyden lisäämistä malliin (Taulukko 12). Herkkyysanalyysien lähtökohtana on päämallina pitämämme Malli 3 Taulukosta 5.

Taulukossa 7 esitellään regressiomallien tulokset erilaisissa tarkasteluvuosiin perustuvissa otoksissa käyttäen Mallin 3 selittäviä muuttujia. Taulukon 7 toisessa sarakkeessa esitellään vertailun helpottamiseksi Mallin 3 tulokset sellaisina, kuin miten ne raportoitiin Taulukossa 5. Taulukon 7 sarakkeissa 3 ja 4 otoksesta poistetaan vuosi 2018 tai vuodet 2018–2019. Sarakkeessa 5 malli sovitetaan pelkkään Valtionkonttorin aineistoon vuosilta 2021–2022 ja sarakkeessa 6 regressiomalli sovitetaan vastaavasti pelkkään Tilastokeskuksen aineistoon (vuodet 2018–2020). Pelkkään Tilastokeskukseen aineistoon perustuvassa mallissa esikoulukokeiluun osallistuneiden osuutta ei voi sisällyttää selittäjäksi, koska vuosina 2018–2020 kokeilu ei ollut käynnissä.

Taulukko 7. Regressiomallin tulokset eri vuosien aineistoilla.

Selitettävä muuttuja: Toimintakulut asukasta kohden					
Vuodet:	2018-22	2019-22	2020-22	2021-22	2018-20
0–5-vuotiaiden osuus	6731.5*** [1723.4]	7386.9*** [1831.9]	8107.0*** [2132.7]	7975.3*** [2509.4]	5922.6*** [1661.3]
6-vuotiaiden osuus	6503.8** [2769.3]	7455.5** [3440.4]	7003.8** [3458.1]	12164.8** [6097.0]	4061.2 [4336.0]
7–12-vuotiaiden osuus	5466.1*** [1753.1]	4986.9** [1992.6]	4905.3** [2097.5]	4724.1* [2508.1]	6218.7*** [1788.5]
13–15-vuotiaiden osuus	9305.9*** [2696.8]	9196.8*** [3037.7]	8580.2** [3355.9]	8613.6** [3752.8]	9002.2*** [2838.1]
Kaksikielinen kunta	36.27 [62.47]	32.65 [64.44]	36.49 [68.53]	57.35 [79.27]	13.64 [66.10]
Vähemmistökielisten osuus	418.2* [232.0]	413.7* [230.1]	373 [235.4]	317.3 [261.8]	496.3* [262.1]
Syntyperä ulkomailla osuus	1202.5** [491.0]	1153.3** [492.8]	1026.3** [499.9]	763.8 [507.9]	1613.1*** [570.3]
Saame kotiseutu kunta	783.7*** [225.5]	804.7*** [241.7]	803.6*** [242.4]	847.3*** [241.8]	733.9*** [217.9]
Syrjäisyysluku	156.8*** [22.79]	151.6*** [23.84]	145.9*** [24.67]	151.3*** [28.49]	160.7*** [22.72]
Saaristoliikenne tai ei yhteyttä osuus	249.0** [101.3]	250.1** [101.9]	243.2** [106.6]	199.5* [119.0]	284.2*** [93.04]
Laskennallinen verotulo/asukas	0.0312* [0.0179]	0.0328* [0.0193]	0.0365* [0.0209]	0.0397 [0.0251]	0.0236 [0.0177]
Väkiluku (tuhansia)	-1.157*** [0.443]	-1.251*** [0.465]	-1.348*** [0.488]	-1.443*** [0.543]	-0.998** [0.429]
Väkiluku(tuhansia) ²	0.00118* [0.000615]	0.00131** [0.000643]	0.00143** [0.000671]	0.00151** [0.000748]	0.00101* [0.000593]
Esikoulukokeiluun osallistujien osuus	6597.3 [4994.6]	6409.7 [4963.6]	6296.6 [4976.0]	6234.8 [5005.3]	
Vakiotermi	559.4*** [85.43]	552.4*** [92.07]	617.7*** [100.7]	582.7*** [117.2]	594.3*** [82.71]
N	1460	1168	875	582	878
R ²	0.641	0.642	0.636	0.621	0.662

Selite: Taulukossa näytetään pns-regressiokertoimet ja hakasuluissa niiden keskivirheet. Tähdet kuvaavat tilastollista merkitsevyyttä (***) $p < 0.01$, ** $p < 0.05$ ja * $p < 0.1$).

Aineiston rajaaminen otosvuosien perusteella osaotoksiin muuttaa estimoitujen parametrien arvoja ja keskivirheitä. 0–5-vuotiaiden osuuden kerroin kasvaa, kun otoksesta poistetaan varhaisimpia vuosia, mutta vastaavasti muiden ikäryhmien kertoimet pienenevät. Muutokset otoksessa eivät kuitenkaan johda keskivirheisiin verrattuna suuriin muutoksiin estimaateissa. Tulosten perusteella ei saada vahvaa näyttöä siitä, että asukaskohtaisten menojen ja selittävien muuttujien yhteys muuttuisi merkittävästi tarkastelujakson aikana. Koko ajanjaksoa 2018–2022 koskevan aineiston käyttäminen vaikuttaa perustellulta.

Taulukossa 8 raportoidaan ikäryhmien perushinnat erilaisissa osaotoksissa. Taulukon toisessa sarakkeessa on vertailun vuoksi Mallin 3 perushinta ja muissa sarakkeissa on erilaisten osaotoksien tuloksista lasketut perushinnat ja perushintojen keskivirheet. Aineiston varhaisten vuosien poistaminen nostaa 0–5-vuotiaiden, 6-vuotiaiden ja 16+ vuotiaiden perushintoja ja vastaavasti laskee 7–12-vuotiaiden ja 13–15-vuotiaiden perushintoja. Osaotoksiin perustuvat perushinnat ovat kuitenkin aina toisessa sarakkeessa esitellyn perushinnan luottamusvälin sisällä (95 prosentin luottamusväli on +/-1,96 keskivirhettä).

Taulukko 8. Ikäryhmien perushinnat eri vuosien aineistoilla (2022 tasossa).

Vuodet:	2018-22	2019-22	2020-22	2021-22	2018-20
0-5 -vuotiaat	7443 [1723]	8099 [1832]	8826 [2134]	8668 [2511]	6538 [1663]
6-vuotiaat	7215 [2775]	8168 [3433]	7722 [3450]	12858 [6087]	4677 [4351]
7-12-vuotiaat	6177 [1757]	5699 [2002]	5624 [2108]	5417 [2519]	6835 [1783]
13-15-vuotiaat	10017 [2657]	9909 [2996]	9299 [3315]	9306 [3713]	9618 [2801]
16+ vuotiaat	711 [70]	712 [72]	719 [72]	693 [76]	616 [72]

Selite: Taulukossa raportoidaan Taulukon 7 regressiomallien tulosten mukaiset ikäryhmien perushinnat, jotka on laskettu kaavalla (2). Perushintojen keskivirheet ovat hakasuluissa.

Taulukossa 9 tarkastellaan regressiomallin tulosten muutoksia, jos malliin lisätään nykyiseen valtiosuusjärjestelmään sisällytettyjä kunnan väestön sosioekonomista asemaa kuvaavia muuttujia (työttömyysaste, ilman tutkintoa 30–54-vuotiaiden osuus ja väestönkasvuaste), joita ei ole sisällytetty tämän raportin päätarkasteluihin. Lisäksi Taulukossa 9 estimoidaan malli ilman kuntakohtaisia kontrollimuuttujia (väkiluku, väkiluvun neliö, laskennallinen verotulo asukasta kohti ja esikoulukokeiluun osallistuneiden osuus) ja taulukon viimeisessä sarakkeessa raportoidaan tulokset, kun regressiomallia on painotettu kunnan osuudella Suomen väestöstä. Laskennalliseen verotuloon sisältyvän mittausvirheen vuoksi Taulukon 9 regressiomallien joukossa on Malli R2a, mihin on sisällytetty laskennallista verotuloa lukuun ottamatta muut kontrollimuuttujat.

Taulukko 9. Vaihtoehtoisten mallien tuloksia.

Selitettävä muuttuja: Toimintakulut asukasta kohden					
Malli	M3	R1	R2a	R2b	R3
0–5-vuotiaiden osuus	6731.5*** [1723.4]	6356.7*** [1850.7]	6531.2*** [1715.2]	5952.9*** [1676.6]	2654.1 [1964.8]
6-vuotiaiden osuus	6503.8** [2769.3]	6225.6** [2802.9]	6554.8** [2770.4]	6789.7** [2781.6]	6495.6 [4575.5]
7–12-vuotiaiden osuus	5466.1*** [1753.1]	4989.6*** [1680.4]	5564.2*** [1748.5]	5876.2*** [1749.1]	9648.3*** [1934.4]
13–15-vuotiaiden osuus	9305.9*** [2696.8]	9304.4*** [2622.0]	9386.5*** [2704.3]	9752.7*** [2769.9]	11269.8*** [3229.9]
Kaksikielinen kunta	36.27 [62.47]	18.34 [65.28]	25.97 [63.69]	13.84 [64.25]	-33.79 [92.06]
Vähemmistökielisten osuus	418.2* [232.0]	451.1* [235.2]	493.2** [232.4]	599.1*** [229.2]	716.7*** [272.5]
Syntyperä ulkomailla osuus	1202.5** [491.0]	1727.9*** [548.7]	1394.9*** [486.2]	818.4* [423.6]	-89.25 [559.0]
Saame kotiseutu kunta	783.7*** [225.5]	786.9*** [229.0]	799.7*** [223.1]	803.1*** [229.0]	553.1*** [143.1]
Syrjäisyysluku	156.8*** [22.79]	174.1*** [24.09]	147.2*** [21.90]	156.2*** [22.04]	202.2*** [34.02]
Saaristoliikenne tai ei yhteyttä osuus	249.0** [101.3]	182.7* [100.9]	264.1*** [99.33]	267.1*** [95.97]	734.2* [423.5]
Työttömyysaste		-780.6* [457.5]			
Ilman tutkintoa 35–50v osuus		-833.1** [399.6]			
Väestönkasvuaste		-1030.9 [970.0]			
Laskennallinen verotulo/asukas	0.0312* [0.0179]	0.0228 [0.0212]			0.0898*** [0.0298]
Väkiluku (tuhansia)	-1.157*** [0.443]	-1.004** [0.493]	-1.020** [0.434]		0.0418 [0.277]
Väkiluku(tuhansia)^2	0.00118* [0.000615]	0.000986 [0.000683]	0.00109* [0.000619]		-0.00011 [0.000418]
Esikoulukokeiluun osallistujien osuus	6597.3 [4994.6]	6001.1 [5005.7]	6173.4 [4957.4]		7008.3 [6560.1]
Vakiotermi	559.4*** [85.43]	809.4*** [156.5]	656.7*** [66.02]	599.4*** [67.47]	220.6 [144.4]
N	1460	1460	1460	1460	1460
R ²	0.641	0.647	0.639	0.635	0.751

Selite: Taulukossa näytetään pns-regressiokertoimet ja hakasuluissa niiden keskivirheet. Tähdet kuvaavat tilastollista merkitsevyyttä (***) $p < 0.01$, ** $p < 0.05$ ja * $p < 0.1$).

Taulukon 9 sarakkeissa R1 sekä R2a ja R2b muutetaan Malliin 3 nähden selittävien muuttujien joukkoa. Kunnan sosioekonomisten tekijöiden lisääminen selittäväksi muuttujiksi sarakkeessa R1 ei näytä merkittävästi muuttavan muiden parametrien kertoimia. Suurimmat muutokset tapahtuvat

kunnan kaksikielisyyttä, 7–12-vuotiaiden osuutta ja saaristoisuutta kuvaavien muuttujien kertoimissa. Estimaatteihin liittyvän tilastollisen epävarmuuden vuoksi mallin R1 kertoimet ovat aina mallin M3 luottamusvälien sisällä. Sosioekonomista asemaa kuvaavat muuttujat eivät näytä olevan yhteydessä korkeampiin asukasta kohti laskettuihin toimintakuluihin. Tulosten perusteella työttömyyden, koulutustaustan ja väestön kasvun poistaminen vos-kriteeristöä vaikuttaisi perustellulta.

Sarakkeessa R2a mallin kontrollimuuttujista poistetaan laskennallinen verotulo asukasta kohden. Muutokset muiden muuttujien kertoimiin verrattuna malliin M3 ovat tällöin vähäisiä. Kaikkien kuntakohtaisten kontrollimuuttujien poistaminen mallista (sarake R2b) näyttää vaikuttavan enemmän estimoitujen parametrien kertoimiin kuin sosioekonomisten tekijöiden lisääminen malliin. Muutoksen taustalla on todennäköisesti se, että muuttujien poistamisen jälkeen poistettujen muuttujien yhteys toimintakuluihin kanavoituu malliin jääneiden kuntien ominaisuuksia kuvaavien muuttujien kautta. Taulukon 9 viimeisessä sarakkeessa (R3) havaintoja painotetaan kunnan väestöosuudella Suomen väestöstä. Painotuksen seurauksena estimoidut kertoimet muuttuvat merkittävästi verrattuna aiempiin muutoksiin ja päämallin tuloksiin. Esimerkiksi 0–5-vuotiaiden osuuden kerroin on yli puolet pienempi kuin Mallin 3 vastaava kerroin. Samalla joidenkin kertoimien etumerkki muuttuu positiivisesta negatiiviseksi. Kaikissa aiemmissa tarkasteluissa kunnan kaksikielisyys lisää kustannuksia, mutta painotetussa mallissa kaksikielisyys näyttää vähentävän kustannuksia. Sama tapahtuu myös syntyperä ulkomailla osuus muuttujan tapauksessa. Tulosten herkkyyks asukasluvulla painotukselle viittaa siihen, että tarkasteltavien vos-kriteereiden vaikutus kustannuksiin on heterogeeninen kuntien välillä. Painottamatonta mallia (M3) voidaan pitää paremmin Suomen kaikkia kuntia edustavana, sillä painotuksessa suuret kaupungit vaikuttavat tuloksiin moninkertaisesti pieniin kuntiin verrattuna.

Taulukossa 10 saaristokuntia kuvaava muuttuja korvataan nykyisen valtiosuusjärjestelmän saaristokuntia kuvaavilla muuttujilla. Sarakkeessa R4 malliin tuodaan selittäjiksi saaristokuntaa kuvaava dummy muuttuja, saaristokuntaa, jonka väestöstä yli 50 prosenttia on ilman kiinteää tieyhteyttä, kuvaava dummy, sekä saaristo-osakuntaa kuvaava dummy. Nykyjärjestelmän perushintoihin verrattuna (Taulukko 1) saaristomuuttujien kertoimet ovat pieniä ja vain valtaosin ilman kiinteää tieyhteyttä olevan saariston muuttuja on tilastollisesti merkitsevä. Taulukon 10 sarakkeessa R5 saaristo-osakuntaa kuvaava dummy muuttuja korvataan saaristo-osakuntien saaristoväestön osuudella. Tämän väestöosuusmuuttujan kerroin on negatiivinen. Mallin muiden selittävien muuttujien kertoimet muuttuvat hyvin vähän verrattuna mallin M3 kertoimiin käytettäessä vaihtoehtoisia saaristoisuutta kuvaavia muuttujia.

Taulukko 10. Vaihtoehtoisten saaristokuntia kuvaavien muuttujien tuloksia.

Selitettävä muuttuja: Toimintakulut asukasta kohden			
Malli	M3	R4	R5
0-5 -vuotiaiden osuus	6731.5*** [1723.4]	6751.7*** [1740.7]	6801.6*** [1734.8]
6-vuotiaiden osuus	6503.8** [2769.3]	6333.4** [2779.1]	6302.6** [2776.6]
7-12-vuotiaiden osuus	5466.1*** [1753.1]	5514.6*** [1776.3]	5473.8*** [1755.0]
13-15-vuotiaiden osuus	9305.9*** [2696.8]	9582.1*** [2682.5]	9355.9*** [2670.2]
Kaksikielinen kunta	36.27 [62.47]	26.36 [61.47]	33.94 [60.98]
Vähemmistökielisten osuus	418.2* [232.0]	403.1* [219.6]	407.3* [223.5]
Syntyperä ulkomailla osuus	1202.5** [491.0]	1222.4** [495.7]	1209.7** [492.9]
Saame kotiseutu kunta	783.7*** [225.5]	790.7*** [225.4]	785.7*** [226.0]
Syrjäisyysluku	156.8*** [22.79]	155.4*** [22.98]	155.5*** [23.02]
Saaristoliikenne tai ei yhteyttä osuus	249.0** [101.3]		
Saaristokunta		93.08 [65.75]	83.73 [65.17]
Saaristo i.k.t. > 50%		171.4*** [24.81]	165.7*** [25.30]
Saaristo-osakunta		26.95 [34.41]	
Saaristoväestön osuus s-osakunnissa			-57.27 [132.8]
Laskennallinen verotulo/asukas	0.0312* [0.0179]	0.0306* [0.0183]	0.0317* [0.0181]
Väkiluku (tuhansia)	-1.157*** [0.443]	-1.182*** [0.434]	-1.167*** [0.445]
Väkiluku(tuhansia)^2	0.00118* [0.000615]	0.00120** [0.000601]	0.00120* [0.000616]
Esikoulukokeiluun osallistujien osuus 2022	6597.3 [4994.6]	6257.8 [4944.7]	5919.3 [4991.3]
Vakiotermi	559.4*** [85.43]	545.9*** [82.98]	554.6*** [84.39]
N	1460	1460	1460
R ²	0.641	0.642	0.642

Selite: Taulukossa näytetään pns-regressiokertoimet ja hakasuluissa niiden keskivirheet. Tähdet kuvaavat tilastollista merkitsevyyttä (***) p<0.01, ** p<0.05 ja * p<0.1).

Taulukossa 11 tarkastellaan vaihtoehtoisia saamelaisten kotiseutualueen mallinnusratkaisuja. Syy tälle tarkastelulle löytyy siitä, että kotiseutualueeseen kuuluu 4 kuntaa (Enontekiö, Inari, Utsjoki ja Sodankylä). Kotiseutualueeseen kuuluvien kuntien vähäisen määrän vuoksi kotiseutualuetta koskevat estimaatit ovat herkkiä regressiomallin määrittelylle ja tämä herkkyys on tulosten läpinäkyvyyden vuoksi syytä tuota esille. Sarakkeessa R6 mallista poistetaan saamelaisten kotiseutualuetta kuvaava muuttuja ja viimeisessä sarakkeessa malliin tuodaan lisäksi kotiseutualuekuntien saamenkielisten osuutta kuvaava muuttuja. Kotiseutualuetta kuvaavan muuttujan poistamisella on yllättävän suuri vaikutus muihin malliin sisällytettyihin kuntien ominaisuuksien parametreihin, kun sarakkeen R6 kertoimia verrataan sarakkeen M3 kertoimiin. Muutoksien kokoluokka on kuitenkin kaikkien selittävien muuttujien kohdalla pienempää kuin estimoidun parametrin keskihajonta. Saamelaisten kotiseutualuetta kuvaavan muuttujan poistaminen mallista johtaa siihen, että kotiseutualueeseen liittyvät kustannukset kanavoituvat muihin mallin kustannuksiin. Sarakkeessa R7 malliin sisällytetään muuttuja kuvaamaan saamelaisten kotiseutualuetta sekä saamenkielisten osuus näissä kunnissa. Tämä mallinnusratkaisu olisi joustava tapa arvioida saamenkielisten palvelujen järjestämisestä aiheutuvia kustannuksia, mutta parametrin estimointi kärsii pienestä saamen kotiseutukuntien lukumäärästä. Tästä syystä saamenkielisten osuuden tuominen malliin tuo verrattain suuria muutoksia mallin muiden kertoimien arvoihin verrattuna malliin M3, jossa on ainoastaan kotiseutualueeseen kuuluvia kuntia kuvaava dummy-muuttuja.

Taulukossa 12 regressiomalleihin lisätään erilaisia asukastiheyttä kuvaavia muuttujia, koska tiivis kuntarakenne voi olla yhteydessä matalampiin asukasta kohti laskettuihin toimintamenoihin. Esimerkiksi koulukyytien järjestäminen voi olla tiheään asutulla alueella halvempaa kuin harvaan asutuilla alueilla. Syrjäisyysluvun avulla pyritään mallintamaan hajanaisuuden vaikutusta kustannuksiin, mutta on syytä testata asukastiheyden selitysvoima syrjäisyyden ohella. Sarakkeessa R8 regressiomalliin lisätään asukastiheys ja tiheyden neliö, sarakkeessa R9 lisätään rajoittamaton asukastiheyskerroin ja viimeisessä sarakkeessa R10 lisätään rajoitettu asukastiheyskerroin. Asukastiheyttä kuvaavilla muuttujilla ei ole suurta vaikutusta muiden malliin sisällytettyjen muuttujien kertoimiin. Asukastiheyden ja tiheyden neliön lisääminen malliin (sarake R8) kasvattaa syntyperä ulkomailla osuuden ja laskennallisen verotulon kertoimia merkittävästi ja väkilukua kuvaavien muuttujien kertoimet pienentyvät verrattuna päätuloksiin (sarake M3). Muut tavat mallintaa asukastiheyttä eivät tuo suuria muutoksia verrattuna malliin M3. Taloudellisesti merkittävä muutos tapahtuu mallissa R9, kun asukastiheyskerroin lisätään selittäjäksi ja tämän seurauksena saamelaisten kotiseutualueen kuntia kuvaavan muuttujan kerroin pienentyy. Muutos liittyy todennäköisesti siihen, että kotiseutukunniksi luokitellut alueet ovat harvaan asuttuja ja tästä syystä asukastiheyden tuominen malliin vaikuttaa Saamen kotiseutukuntien kertoimeen.

Taulukko 11. Vaihtoehtoisten saamelaisten kotiseutualueen kuntia kuvaavien muuttujien tuloksia.

Selitettävä muuttuja: Toimintakulut asukasta kohden			
Malli	M3	R6	R7
0–5-vuotiaiden osuus	6731.5*** [1723.4]	7088.6*** [1967.0]	6308.0*** [1672.8]
6-vuotiaiden osuus	6503.8** [2769.3]	6032.4** [2974.6]	7387.2*** [2737.2]
7–12-vuotiaiden osuus	5466.1*** [1753.1]	6012.9*** [2129.4]	6356.1*** [1525.0]
13–15-vuotiaiden osuus	9305.9*** [2696.8]	8674.5** [3534.6]	7320.8*** [2187.7]
Kaksikielinen kunta	36.27 [62.47]	17.27 [69.47]	50.16 [61.15]
Vähemmistökielisten osuus	418.2* [232.0]	395.2 [247.1]	398.9* [228.3]
Syntyperä ulkomailla osuus	1202.5** [491.0]	1469.7** [682.2]	905.7** [427.6]
Saame kotiseutu kunta	783.7*** [225.5]		357.9*** [73.19]
Saamenkielisten osuus			2765.5*** [169.4]
Syrjäisyysluku	156.8*** [22.79]	207.3*** [37.94]	152.4*** [22.65]
Saaristoliikenne tai ei yhteyttä osuus	249.0** [101.3]	209.0** [103.9]	239.6** [98.53]
Laskennallinen verotulo/asukas	0.0312* [0.0179]	0.0555*** [0.0208]	0.0353** [0.0176]
Väkiluku (tuhansia)	-1.157*** [0.443]	-1.276*** [0.487]	-1.060** [0.429]
Väkiluku(tuhansia) ²	0.00118* [0.000615]	0.00120* [0.000640]	0.00107* [0.000599]
Esikoulukokeiluun osallistujien osuus 2022	6597.3 [4994.6]	8533.1 [6430.5]	4145.8 [4766.5]
Vakiotermi	559.4*** [85.43]	432.4*** [123.3]	576.3*** [84.38]
N	1460	1460	1460
R ²	0.641	0.573	0.666

Selite: Taulukossa näytetään pns-regressiokertoimet ja hakasuluissa niiden keskivirheet. Tähdet kuvaavat tilastollista merkitsevyyttä (***) p<0.01, ** p<0.05 ja * p<0.1).

Taulukko 12. Asukastiheyttä kuvaavia muuttujia lisätty malliin.

Selitettävä muuttuja: Toimintakulut asukasta kohden				
Malli	M3	R8	R9	R10
0–5-vuotiaiden osuus	6731.5*** [1723.4]	7061.2*** [1716.8]	6406.7*** [1687.9]	6663.0*** [1659.2]
6-vuotiaiden osuus	6503.8** [2769.3]	6815.3** [2744.4]	6460.1** [2735.8]	6429.5** [2742.2]
7–12-vuotiaiden osuus	5466.1*** [1753.1]	5315.6*** [1746.5]	5589.1*** [1714.2]	5384.3*** [1744.2]
13–15-vuotiaiden osuus	9305.9*** [2696.8]	9682.4*** [2714.1]	9575.0*** [2655.6]	9487.3*** [2702.4]
Kaksikielinen kunta	36.27 [62.47]	28.25 [56.23]	32.7 [62.26]	30.81 [62.69]
Vähemmistökielisten osuus	418.2* [232.0]	364.8* [215.7]	442.2* [227.9]	449.1** [226.9]
Syntyperä ulkomailla osuus	1202.5** [491.0]	1837.6*** [530.6]	1290.2** [499.6]	1298.6** [505.8]
Saame kotiseutu kunta	783.7*** [225.5]	756.4*** [221.2]	652.1*** [199.9]	710.1*** [239.8]
Syrjäisyysluku	156.8*** [22.79]	168.8*** [23.60]	134.0*** [25.27]	122.5*** [32.21]
Saaristoliikenne tai ei yhteyttä osuus	249.0** [101.3]	241.6** [100.5]	265.4** [102.5]	264.7*** [101.8]
Asukastiheys		-0.450*** [0.157]		
Asukastiheys ²		0.000177 [0.000112]		
Asukastiheys kerroin			3.13 [2.153]	
Asukastiheys kerroin (max 20)				6.775 [6.081]
Laskennallinen verotulo/asukas	0.0312* [0.0179]	0.0591** [0.0235]	0.0245 [0.0184]	0.0273 [0.0171]
Väkiluku (tuhansia)	-1.157*** [0.443]	-0.463 [0.643]	-1.136** [0.441]	-1.156*** [0.444]
Väkiluku(tuhansia) ²	0.00118* [0.000615]	-0.00076 [0.00226]	0.00118* [0.000615]	0.00119* [0.000618]
Esikoulukokeiluun osallistujien osuus 2022	6597.3 [4994.6]	6596.6 [4931.0]	6070.3 [4860.3]	6114.4 [4942.8]
Vakiotermi	559.4*** [85.43]	430.0*** [107.4]	578.4*** [84.12]	569.8*** [82.33]
N	1460	1460	1460	1460
R ²	0.641	0.648	0.646	0.643

Selite: Taulukossa näytetään pns-regressiokertoimet ja hakasuluissa niiden keskivirheet. Tähdet kuvaavat tilastollista merkitsevyyttä (***) p<0.01, ** p<0.05 ja * p<0.1).

Taulukossa 13 raportoidaan ikäryhmien perushinnat ja perushintojen keskivirheet Taulukoissa 9–12 esitellyistä vaihtoehtoisista regressiomalleista. Tulosten vertailun helpottamiseksi taulukko sisältää myös päätuloksiksi valitusta Malli 3:sta lasketut perushinnat. Suurin muutos verrattuna Malli 3:n perushintoihin tapahtuu mallissa R3, jossa regressiomallia painotetaan kunnan väestöosuudella. Painotuksen seurauksena esimerkiksi 0–5-vuotiaiden perushinta laskee noin 7400 eurosta noin 3200 euroon ja 7–12-vuotiaiden perushinta kasvaa noin 4000 eurolla. Ikäryhmien perushinnat eivät pääosin ole kovin herkkiä vaihtoehtoisille selittäville muuttujille (Mallit R1–R2b ja R4–10), kun huomioidaan perushinnoille lasketut keskivirheet. Mallin R7 erot ovat merkittävimmät Malliin M3 verrattuna. 7–12-vuotiaiden perushinta kasvaa noin 1000 eurolla Mallin R7 kohdalla, kun saamelaisten kotiseutualuetta mallinnetaan myös saamenkielisten väestöosuuksilla kyseisissä kunnissa. Myös 13–15-vuotiaiden perushinta muuttuu merkittävästi Mallissa R7 Malliin 3 verrattuna. 16+ ryhmän perushinta kasvaa merkittävästi Mallin R1 kohdalla, kun selittäviksi muuttujiksi on lisätty kunnan sosioekonomista asemaa kuvaavia muuttujia.

Taulukko 13. Ikäryhmien perushinnat vaihtoehtoisissa malleissa (2022 taso).

Malli	M3	R1	R2a	R2b	R3	R4	R5
Paneeli A							
0–5-vuotiaat	7443 [1723]	7292 [1843]	7241 [1716]	6678 [1678]	3184 [1983]	7448 [1742]	7511 [1738]
6-vuotiaat	7215 [2775]	7161 [2818]	7264 [2775]	7515 [2785]	7026 [4577]	7029 [2783]	7012 [2781]
7–12-vuotiaat	6177 [1757]	5925 [1669]	6274 [1751]	6601 [1750]	10179 [1895]	6210 [1778]	6183 [1755]
13–15-vuotiaat	10017 [2657]	10239 [2577]	10096 [2664]	10478 [2731]	11800 [3187]	10278 [2646]	10065 [2632]
16+ vuotiaat	711 [70]	935 [125]	710 [71]	725 [70]	530 [115]	696 [68]	709 [70]
Malli	M3	R6	R7	R8	R9	R10	
Paneeli B							
0–5-vuotiaat	7443 [1723]	7746 [1936]	7057 [1679]	7737 [1715]	7114 [1692]	7371 [1660]	
6-vuotiaat	7215 [2775]	6690 [3003]	8136 [2739]	7491 [2750]	7168 [2742]	7138 [2750]	
7–12-vuotiaat	6177 [1757]	6670 [2167]	7105 [1515]	5992 [1752]	6297 [1714]	6093 [1748]	
13–15-vuotiaat	10017 [2657]	9332 [3464]	8070 [2158]	10359 [2672]	10283 [2617]	10196 [2662]	
16+ vuotiaat	711 [70]	657 [99]	749 [64]	676 [72]	708 [69]	708 [70]	

Selite: Taulukossa raportoidaan Taulukoiden 9–12 regressiomallien tulosten mukaiset ikäryhmien perushinnat, jotka on laskettu kaavalla (2). Perushintojen keskivirheet ovat hakasuluissa.

Taulukon 13 perusteella perushinnoissa suurimmat muutokset aiheutti regressiomallin painottaminen kunnan väestöosuudella Mallissa R3 ja saamenkielisten väestöosuuden lisääminen malliin selittäjäksi mallissa R7. On syytä huomata, että perushinnan muutokset eivät aina ole

samansuuruisia. Joillekin ikäryhmille regressiomallin painottaminen voi johtaa pienentyneeseen perushintaan, mutta samalla toisen ikäryhmän perushinta voi kasvaa.

5. Tulosten hyödyntäminen valtionosuusjärjestelmän uudistamisessa

Luvussa 4 esitellyt tulokset tarjoavat hyvän lähtökohdan valtionosuusjärjestelmän uudistamiseen. Regressioanalyysi soveltuu hyvin perushintojen ja kriteerien kertoimien suuruuden estimointiin, koska kaikkien kriteerien kertoimet määräytyvät samanaikaisesti siten, että malli sopii tutkittavaan aineistoon mahdollisimman hyvin. Regressioanalyysit perustuvat nykyiseen valtionosuusjärjestelmään, mutta tarkasteluihin sisällytettäviä kriteerejä on pyritty muokkaamaan siten, että ne kuvaisivat aiempaa tarkemmin palvelujen kustannuksia aiheuttavia tekijöitä. Osa aiemmin valtionosuusjärjestelmään kuuluneista kriteereistä, joiden teoreettinen perusta on heikko, on poistettu kokonaan tai siirretty raportin herkkystarkasteluihin.

Regressiomallin tuloksiin liittyy mahdollisia harhan lähteitä ja toisinaan suurta tilastollista epävarmuutta, mutta tuloksia voidaan silti pitää parhaana käytettävissä olevana arviona kuntien ominaisuuksien ja toimintamenojen välisistä riippuvuussuhteista. Tästä lähtökohdasta huolimatta tuloksia ei ole syytä hyödyntää valtionosuusjärjestelmän uudistamisessa suoraan ilman harkintaa. Merkittävin tuloksiin liittyvä haaste on 16+ ikäryhmän perushinnan kokoluokka. Tilastomalleihin perustuva perushinta on yli kymmenkertainen verrattuna nykyisen järjestelmän vastaavaan perushintaan ja perushinta on moninkertainen myös tilastoista saatavaan vastaavaan keskimääräiseen kustannukseen.

Valtionosuusjärjestelmän kehittämisessä 16+ ikäryhmän perushinnan määrittämisessä on syytä käyttää regressioanalyysin tulosten lisäksi muuta tilastotietoa ja asiantuntijaharkintaa. Yksi mahdollinen tapa on yhdistää tilastoihin suoraan perustuvat toimintamenot ja luvussa 4 esitellyt regressiomallien tulokset. Tällöin 16+ ryhmän perushinta perustuisi esimerkiksi historiallisiin kyseisen ryhmän kustannuksiin ja muut perushinnat ja kriteerien kertoimet perustuisivat regressiotuloksiin. Asetettaessa 16+ ryhmän perushinta regressioanalyysin tuloksia alemmaksi on samalla perusteltua korottaa muiden ikäryhmien perushintoja (esim. samalla rahasummalla). Tämä on yksi mahdollinen tapa huomioida regressiomalleihin perustuvan lähestymistavan suurin haaste, mutta on muitakin tapoja sisällyttää valtionosuuslaskelmiin regressiotuloksien ulkopuolista kustannustietoa. Selvää on kuitenkin se, että tässä raportissa estimoitu 16+ ryhmän perushinta on epärealistisen suuri verrattuna tarjolla oleviin verrokkietoihin ja nykyiseen vos-järjestelmän arvoon.

Lisäksi joidenkin valtionosuuskriteerien kertoimet poikkeavat merkittävästi nykyisen järjestelmän kertoimista (esimerkiksi saaristoisuuden tapauksessa) ja järjestelmän päivittäminen vastaamaan luvussa 4 esiteltyjä kertoimia johtaa joidenkin kuntien kohdalla merkittäviin muutoksiin valtionosuusjärjestelmän kautta kanavoituvassa rahoituksessa. Osa kunnista saisi tuloksia sovellettaessa merkittävää lisärahoitusta, mutta vastaavasti osa kunnista menettäisi rahoitusta. Mikäli muutoksia halutaan tasoittaa, voidaan tässä hyödyntää valtionosuusjärjestelmään tuotavia lisäosia tai muita erillisratkaisuja. Myös siirtymäaika helpottaa kuntien sopeutumista uuteen valtionosuuksien tasoon.

6. Lopuksi

Tässä tutkimuksessa käytetään eri lähteistä koottua aineistoa kuntien menoista ja niitä mahdollisesti selittävästä ominaisuuksista vuosilta 2018–2022. Kuntien ominaisuuksien yhteyttä kustannuksiin tutkitaan regressiomallilla, jossa sote-uudistuksen jälkeisiä palveluja vastaavia toimintamenoja selitetään mahdollisilla valtionosuuskriteereillä. Tulokset eivät välttämättä anna luotettavaa kuvaa menojen ja kuntien ominaisuuksien välisestä syy-seuraus suhteesta, sillä tutkittavien muuttujien kautta saattaa kanavoitua mallin ulkopuolisten tekijöiden vaikutuksia ja tuloksiin liittyy usein myös suurta tilastollista epävarmuutta varsin pienestä otoskoosta johtuen. Regressioanalyysin tuloksia voidaan silti pitää hyvänä lähtökohtana nykyisen valtionosuusjärjestelmän uudistamisessa. Kerroinestimaatit kuvaavat selittävien muuttujien keskimääräistä yhteyttä tutkittavien palveluiden kustannuksiin. Ne tarjoavat objektiivisen havaintoaineistoon perustuvan vertailukohtan nykyisen valtionosuusjärjestelmän kertoimille, jotka perustuvat vanhentuneisiin selvityksiin sekä asiantuntijoiden ja poliitikkojen näkemyksiin.

Merkittävä haaste valtionosuusjärjestelmän perushintojen estimoinnissa on kustannusaineistojen saatavuus. Tällä hetkellä kuntien kustannusaineistot ovat jakautuneet Tilastokeskuksen ja Valtiokonttorin palveluihin. Aineistojen lataaminen kahdesta palvelusta on epäkäytännöllistä, ja määritelmässä on eroja. Tulevaisuuden tarkastelut voivat perustua pelkästään Valtionkonttorin aineistoihin, kun palveluun kertyy lisää vuosia. Tällöin kunta-aineistojen latauspalvelussa tulee ottaa huomioon tutkimuskäytön näkökulma, siten että kattavia aineistoja voi ladata kerralla. Kunkin vuoden kuntajaon mukaisten aineistojen saatavuus on myös tärkeää tutkimuksen kannalta.

Lähteet

- Harjunen, O. ja Saarimaa, T. (2012): Ikärakenne kuntien valtionosuusjärjestelmässä: laskennalliset kustannukset sosiaali- ja terveydenhuollossa sekä esi- ja perusopetuksessa. VATT-muistiot, 23.
- Häkkinen, U., Nguyen, L., Pekurinen, M. ja Peltola, M. (2009): Tutkimus terveyden- ja vanhustenhuollon tarve- ja valtionosuuskriteereistä. Jyväskylä.
- Lehtonen, S., Lyytikäinen, T. ja Moisio, A. (2008a): Kuntien valtionosuuskriteerit tarkastelussa: esi- ja perusopetus, päivähoito, kirjastot ja kulttuuritoimi. VATT keskustelualoitteita 451.
- Lehtonen, S., Lyytikäinen, T. ja Moisio, A. (2008b): Kuntien rahoitus- ja valtionosuusjärjestelmä: vaihtoehtoja uudistuksen toteuttamiseksi. VATT tutkimuksia 141.

Liite A: Muuttujakuvaukset

Tässä liiteluvussa kuvataan, kuinka tilastollisiin malleihin sisällytetyt muuttujat on muodostettu ja mistä aineistolähteestä tiedot ovat peräisin.³

Toimintamenot

Kunnan väestöön suhteutettuja toimintamenoja käytetään selitettävänä muuttujana kaikissa tilastomalleissa. Tarkasteluissa käytettävät toimintameno tiedot ovat peräisin Tilastokeskuksen ja Valtionkonttorin rekisteristä. Aineistoon sisällytetään vain Manner-Suomen kunnat. Taulukossa L1 luetellaan tarkasteluun sisällytetyt toimintamenukat. Tilastokeskuksen tiedot kattavat vuoden 2018–2020 ja tietolähteenä on käytetty Tilastokeskuksen Kuntien käyttötalous tilastoa⁴. Tilastosta on haettu kaikkien Manner-Suomen kuntien palveluluokkakohtaiset toimintamenot. Valtionkonttorin tiedot kattavat vuodet 2021 ja 2022 ja Valtionkonttorista haetut tiedot perustuvat palveluluokkakohtaisiin Kunnan ja kuntayhtymien tilinpäätöstietoihin⁵. Valtionkonttorin rekisteristä on poimittu kuntakohtaisesti Manner-Suomen kuntien tiedot toimintakuluista. Valtion konttorin palvelussa hakukoneen sektoreina on käytetty luokittelua sisäiset ja ulkoiset menot. Tilastokeskuksen ja Valtionkonttorin rekisterien yhdistäminen kunnanumeron perusteella tuottaa vuodet 2018–2022 kattavan paneeliaineiston kuntien toimintamenoista.

Taulukko L1. Toimintakulukategoriat Tilastokeskuksen ja Valtionkonttorin aineistoissa.

<u>Tilastokeskus (2018–2020)</u>	<u>Valtionkonttori (2021–2022)</u>
Ympäristöterveydenhuolto	Elintarvikevalvonta ja neuvonta
Kirjastotoiminta	Eläinten terveyden ja hyvinvoinnin valvonta
Taiteen perusopetus	Eläinlääkäripalvelut
Muu kulttuuritoiminta	Esiopetus
Perusopetus	Kirjastopalvelut
Esiopetus	Muu varhaiskasvatus
Varhaiskasvatus	Perhepäivähoito
	Perusopetus
	Päiväkotitoiminta
	Taiteen perusopetus + tukeva toiminta
	Terveydensuojelun edistäminen ja valvonta
	Tupakkalain valvonta
	Yleisen kulttuuritoiminnan palvelut

³ Tutkimuksen aineiston muodostamisessa ja analysoinnissa käytetyt koodit sekä tutkimusaineisto on julkaistu tämän tutkimuksen yhteydessä Suomen kansalliskirjaston ylläpitämässä Doria julkaisuarkistossa.

⁴ https://pxdata.stat.fi/PxWeb/pxweb/fi/Kuntien_talous_ ja_toiminta/Kuntien_talous_ ja_toiminta Kunnat/

⁵ <https://www.tutkihallintoa.fi/kunnat/kuntien-ja-kuntayhtymien-talous/kunnan-ja-kuntayhtymien-tilinpaatostiedot-palveluluokkakohtaisesti/>

Selite: Taulukossa luetellaan tutkimuksen kustannusmuuttujaan sisällytetyt toimintamenokategoriat Tilastokeskuksen ja Valtiokonttorin aineistoissa.

Ennen varsinaista tilastoanalyysiä kuntakohtaisia toimintakuluja muokataan viiden seikan suhteen. Tehtävät muutokset ovat seuraavat: i) lisätään kuntien menolajikohtaiset poistot menoihin ii) toimintamenoista vähennetään saadut kotikuntakorvaukset iii) toimintamenoista vähennetään kotihoidontuen kuntalisä, iv) vuosien 2018–2020 ajalta perusopetuksen kustannuksista poistetaan aamu- ja iltapäivätoiminnan kustannukset ja v) kustannukset deflatoidaan Tilastokeskuksen opetus- ja kulttuuritoimen hintakehitystä kuvaavalla julkisten menojen hintaindeksillä vuoden 2022 arvoon. Poistot lisätään kustannuksiin, koska tällöin toimintakulumuuttuja kohtelee yhtäläisellä tavalla omien tilojen ja laitteiden käyttöä ns. leasing kustannuksiin verrattuna. Saadut kotikuntakorvaukset vähennetään toimintamenoista, koska tällöin kustannukset vastaavat paremmin kunnan oman väestön palvelukäytön kustannuksia. Tieto kotikuntakorvauksista perustuu Kelan tietoihin. Kotihoidontuen kuntalisä vähennetään kustannuksista, koska kyseisen tuki ei kuulu valtionosuuksien kustannuspohjaan. Samalla perusteella vuosien 2018–2020 aineistosta poistetaan arvio perusopetuksen kustannuksiin sisältyvät aamu- ja iltapäivätoiminnan kustannukset. Vuosien 2021–2022 aineistossa aamu- ja iltapäivätoiminnan kustannukset on erotettu omaksi eräkseen, joten näiltä vuosilta niiden poistaminen on suoraviivaista. Aamu- iltapäivätoiminnan kustannukset poistetaan vuosina 2018–2020 laskemalla vuoden 2022 aineistossa perusopetuksen osuus perusopetuksen ja aamu- ja iltapäivätoiminnan kustannuksien summasta kuntakohtaisesti. Tällä osuudella korjataan kuntakohtaisia perusopetuksen kustannuksia vuosina 2018–2020 kertomalla perusopetuksen kustannukset vuoden 2022 kustannusosuudella. Kuntaliitosten vuoksi muutama vuosien 2018–2020 aikana toiminut kunta on lopettanut toimintansa ja näille kunnille käytetään kustannusosuutena muiden kuntien kustannusosuuden mediaania. Toimintamenot deflatoidaan opetus - ja kulttuuritoimen hintakehitystä kuvaavalla hintaindeksillä vuoden 2022 arvoon, koska suurin osa tarkasteluun sisällytetyistä kustannuskategorioista liittyy kuntien opetus - tai kulttuuritoimeen.

Tilastollisissa analyyseissä käytettävä toimintakulumuuttuja muodostetaan laskemalla yhteen kaikki tarkasteluun sisällytettävät kustannuskategorian jokaisessa kunnassa jokaisena tarkasteluvuotena. Tämän jälkeen summattu toimintakulumuuttuja jaetaan kunnan kokonaisväestöllä. Tämä muuttuja kuvaa kuinka suuri määrä kunnan toimintamenoista kohdistuu yksittäistä kuntalaista kohti. Raportin kuvailevissa tarkasteluissa toimintakulut summataan kategorioihin varhaiskasvatus, esiopetus, perusopetus ja muut toimintakulut. Varhaiskasvatus, esiopetus ja perusopetus suhteutetaan palveluiden piirissä laskennallisesti olevaan kunnan väestöön ja muut kustannukset suhteutetaan kunnan kokonaisväestöön. Tämä jaottelu auttaa hahmottamaan yksittäisten kustannuserien taloudellista merkittävyyttä.

Muut muuttujat

Ikäryhmät

Tilastollisissa tarkasteluissa hyödynnetään tietoa kunnan ikärakenteesta. Ikärakennemuuttujat muodostetaan seuraavien jaottelujen avulla: 0–5-vuotiaat, 6-vuotiaat, 7–12-vuotiaat, 13–15-vuotiaat ja yli 15-vuotiaat. Jokaiseen ryhmään sisältyvän väestön määrä lasketaan yhteen ja jaetaan kunnan kokonaisväestöllä. Kuntien väestötiedot perustuvat Tilastokeskuksen väestörakennetilastoon, joka kuvaa kuntien väestöä vuoden lopussa. Tilastollisiin malleihin sisällytetään selittäviksi muuttujiksi kaikki muodostetut ikäryhmät paitsi yli 15-vuotiaiden ryhmä. Ikäryhmien tulkinnassa on hyvä

huomata se, että kaikki ikäryhmän jäsenet eivät käytä kunnan tuottamia palveluja. Esimerkiksi varhaiskasvatuksen tapauksessa kaikki 0–5-vuotiaat eivät käytä kunnan tuottamia palveluita ja tästä syntyy mittavirhettä.

Laskennallinen verotulo / asukas

Kaikkiin regressiomalleihin sisällytetään kontrollimuuttujaksi ikäryhmien ja vuosimuuttujien lisäksi kunnan laskennallinen asukkaisiin suhteutettu verotulo. Tämä muuttuja auttaa vakioimaan kuntien välisiä eroja tuotettujen palveluiden laadussa. Esimerkiksi koulujen luokkakoot voivat olla pienempiä tai varhaiskasvatuksen palvelukapasiteetti on riittävän suuri kunnan asukkaille. Varakkaat kunnat voivat pystyä tuottamaan lakisääteisiä palveluita laadukkaammin kuin pienen laskennallisen verotulon kunnat ja tästä voi seurata se, että laskennallinen verotulo voi vaikuttaa kotitalouksien sijaintipäätöksiin. Laskennallinen verotulo on poimittu valtiovarainministeriön julkaisemista valtionosuuslaskelmista. Laskennallinen verotulo ei mittaa kunnan tulopohjaa täysimääräisesti, sillä kiinteistöveron veropohja sisältyy laskennalliseen verotuloon vain 50-prosenttisesti ja voimalaitosten kiinteistöveron osalta huomioidaan vain ydinvoimalaitoksista maksettava kiinteistövero.

Kunnan kokonaisväestö

Kaikkiin tilastollisiin malleihin lisätään selittäviksi muuttujiksi myös kunnan kokonaisväkiluku ja kokonaisväkiluvun neliö. Tämä mallinnusratkaisu ottaa huomioon sen, että palvelutuotannossa voi olla kunnan väestöön liittyviä ns. skaalaetuja, jolloin palvelutuotannon määrän kasvaessa kustannukset tuotettua palvelua kohti voivat muuttua. Kuntien väestötiedot perustuvat Tilastokeskuksen väestörakennetilastoon, joka kuvaa kuntien väestöä vuoden lopussa.

Kunnan asukastiheyttä kuvaavat muuttujat

Päätulosten asukastiheyteen liittyvissä herkkyystarkasteluissa (Taulukko 12) vaihtoehtoihin malleihin tuodaan asukastiheys ja asukastiheyden neliö, rajoittamaton asukastiheyskerroin ja rajattu asukastiheyskerroin. Näiden herkkyystarkastelujen tarkoitus on tuoda esille kuinka herkkiä tutkimuksen päätulokset ovat asukastiheyden vaihtoehtoisille mallinnusratkaisuille. Tutkimuksessa käytettävät asukastiheystiedot on poimittu Valtionvarainministeriön julkaisemista valtionosuuslaskemista.

Esikoulukokeiluun osallistuvien osuus

Kaikkiin tilastollisiin malleihin sisällytetään selittäväksi muuttujaksi myös esikoulukokeiluun osallistuvien osuus kunnassa vuonna 2022. Esikoulukokeilu on voinut vaikuttaa varhaiskasvatuksen ja esikoulun järjestämisen toimintameneihin ja tästä syystä kokeilun piirissä olevien henkilöiden osuuden yhteys toimintameneihin kontrolloidaan. Kokeilun huomioiminen auttaa 0–5- vuotiaiden ikäryhmän ja 6-vuotiaiden ikäryhmän kertoimien identifioinnissa, koska kokeilun huomiotta jättäminen voisi heijastaa kokeiluun liittyviä kustannuksia kiinnostuksen kohteena oleviin ikäryhmämuuttujiin. Muuttuja on saatu VATT:n tutkimusryhmältä.

Kieliasemaa kuvaavat muuttujat

Tilastollisiin malleihin sisällytetään kunnan kieliasemaa kuvaavista tiedoista tieto kunnan kaksikielisyydestä, vähemmistökielisten osuudesta, kuuluuko kunta saamelaiden kotiseutukuntiin, saamenkielisten osuus kotiseutukunnassa sekä tieto kunnan asukkaiden syntyperästä. Kieliasemaa kuvaavat muuttujat sisällytetään tarkasteluihin, koska kuntien palvelutuotanto voi muodostua kalliimmaksi, mikäli palveluja tulee järjestää usealla kielellä. Kaksikielisyyttä kuvaava muuttuja erottelee kunnat sen perusteella, onko kunta kaksikielinen vai ei. Samalla tavalla kunnat luokitellaan

saamelaisten kotiseutukuntiin ja kotiseutukuntien ulkopuolisiin kuntiin. Kaksikielisyyttä, vähemmistökielisten osuutta, saamelaisten kotiseutukuntaa ja saamenkielisten osuus kotiseutukunnassa kuvaavat muuttajat on poimittu valtiovarainministeriön julkaisemista valtiosuosuuskalkelmista.

Kunnan asukkaiden syntyperä muuttujassa taustaltaan ulkomaalaisiksi määritellään henkilöt, joiden vanhemmat (tai ainoa vanhempi) ovat syntyneet ulkomailla. Muuttuja perustuu Tilastokeskuksen Syntyperä ja taustamaa sukupuolen mukaan kunnittain tilastoon⁶.

Aiemmissa valtiosuosuuskalkelmissa kunnan vieraskielisten määrää on käytetty yhtenä kieliasemaa kuvaavana muuttujana. Vieraskielisten määrän käyttäminen on ongelmallista, koska muuttuja ei perustu mitattuun kielitaitoon vaan tieto on yksilön itse ilmoittama. On mahdollista, että suomalaisen peruskoulun suorittanut henkilö voi ilmoittaa äidinkielenkseen Suomen virallisista kielistä poikkeavan kielen, jolloin henkilö määritellään tilastoissa vieraskieliseksi. Tämä muuttuja on nyt korvattu kunnan asukkaiden syntyperää kuvaavalla muuttujalla, joka voi soveltua paremmin vieraskielisyydestä aiheutuvien lisäkustannusten mallintamiseen.

Syrjäisyys ja saaristoisuus

Tilastollisiin malleihin sisällytetään myös kunnan syrjäisyyttä kuvaavia tekijöitä. Näiden muuttujien tarkoitus on vakioida kuntien välisiä eroja toimintakuluissa, jotka ovat peräisin eroista kuntien syrjäisyydessä. Kunnan syrjäisyyttä mallinnetaan tie-etäisyyteen perustuvan syrjäisyysluvun ja saaristoliikenteen tai ilman pysyvää liikenneyhteyttä olevan väestön osuus kunnan kokonaisväestöstä. Saaristoisuusmuuttuja perustuu Suomen Ympäristökeskuksen Liiteri tietokantaan ja on saatavilla vain vuosille 2021 ja 2022. Muille vuosille käytetään kunnan vuoden 2021 lukua. Syrjäisyysluku perustuu Tilastokeskuksen toimittamiin tietoihin. Tiedot olivat saatavilla vuosilta 2022, 2021 ja 2019. Vuoden 2020 osalta käytetään vuoden 2021 syrjäisyyslukua ja vuoden 2018 puuttuva tieto korvataan vuoden 2019 luvulla.

Syrjäisyysluvun muodostaminen on kuvattu laissa seuraavasti: ”Syrjäisyysluku muodostuu kahden väestöpohjaindeksin summasta, joiden kummankin minimiarvo on 0. Paikallinen väestöpohjaindeksi mittaa väestön määrää 25 kilometrin etäisyydellä ja seudullinen väestöpohjaindeksi 50 kilometrin etäisyydellä kunnan väestöllisestä keskipisteestä. Kunnan väestöllinen keskipiste on kunnan laskennallinen väestön keskipiste, joka lasketaan pinta-alaltaan yhden neliökilometrin suuruista ruuduista käyttäen painomuuttujina kunkin asutun ruudun väestömäärää ja sijaintimuuttujana ruudun keskipisteen koordinaatteja.

Paikallinen väestöpohjaindeksi saadaan vähentämällä 15 000:sta enintään 25 kilometrin etäisyydellä asuvan väestön määrä ja jakamalla erotus 15 000:lla. Seudullinen väestöpohjaindeksi saadaan vähentämällä 60 000:sta enintään 50 kilometrin etäisyydellä asuvan väestön määrä ja jakamalla erotus 60 000:lla. Indeksit lasketaan Tilastokeskuksen viiden vuoden välein tarkistamasta tilastosta, joka koskee väestön jakautumista 62 500 neliömetrin kokoisille alueille. Laskennassa käytettävät etäisyydet määritetään tieverkossa.”

Syrjäisyysluvun laskennassa tieverkon jatkeena olevat lautta- ja lossireitit katsotaan osaksi tieverkkoa. Niistä aiheutuvaa viivästystä matka-aikoihin ei huomioida. Siten syrjäisyysluku saattaa aliarvioida saaristokuntien syrjäisyyttä.

⁶ https://pxdata.stat.fi/PxWeb/pxweb/fi/StatFin/StatFin__vaerak/statfin_vaerak_pxt_11rv.px/

Sosioekonominen asema

Tulosten herkkyystarkasteluissa malleihin lisätään vielä kunnan sosioekonomista asemaa kuvaavia muuttujia. Kuntien sosioekonomisista tekijöistä herkkyyssanalyseissä huomioidaan kunnan työttömyysaste, väestön kasvuaste ja ilman tutkintoa olevien 30–54-vuotiaiden osuus. Näiden muuttujien sisällyttämistä valtionosuusjärjestelmään voidaan pitää teoreettisesti heikosti perusteltuna, sillä ei ole selvää miksi ja minkä suuntaisesti ne vaikuttaisivat tutkittaviin toimintamenoihin. Muuttujia on silti testattu osana mallia, sillä ne ovat mukana nykyisessä vos-järjestelmässä. Kunnan sosioekonomista asemaa koskevat tiedot on poimittu Valtiovarainministeriön julkaisemista valtionosuuslaskelmista, ja ne perustuvat Tilastokeskuksen ja TEM:n tilastoihin.