

Työvoimatutkimuksen uudistus 2021 ja sen vaikutukset aikasarjoihin

Tiivistelmä

Vuoden 2021 alussa työvoimatutkimuksen tietosisältöä, tiedonkeruuta ja estimointimenetelmää uudistettiin. Uudistus ei liittynyt koronaviruspandemiaan, vaan yleisesti työvoimatutkimuksen kehittämiseen.

Työvoimatutkimuksen kyselylomakkeen sisältöä muutettiin uuden EU-lainsäädännön mukaiseksi. Keskeisimmät muutokset koskivat työaikatietoja sekä työllisen määritelmää.

Tiedonkeruun uudistuksessa otettiin käyttöön verkkotiedonkeruu haastattelujen rinnalle.

Uudistetussa estimointimenetelmässä hyödynnetään entistä laajemmin rekisterilähteistä saatavaa lisäinformaatiota, joka auttaa oikaisemaan vastauskadon aiheuttamaa vinoumaa ja kaventamaan estimaatin virhemarginaalia.

Vanhaan menetelmään perustuvat luvut eivät ole vertailukelpoisia uudella menetelmällä tuotettujen kanssa.

Miten ja miksi uudistus tehtiin

Työvoimatutkimuksen kokonaisuutta uudistettiin vuoden 2021 alusta lähtien uuden EU-lainsäädännön mukaiseksi, ja samalla toteutettiin merkittäviä kansallisia uudistuksia tiedonkeruuseen ja estimointiin. Uudistusten tavoitteena on helpottaa tutkimukseen vastaamista sekä parantaa tulosten laatua ja kansainvälistä vertailtavuutta.

Työvoimatutkimus on otostutkimus, jossa tiedot kerätään kysymällä niitä vastaajilta. Vuoden 2020 loppuun asti tietoja kerättiin puhelinhaastatteluin ja osin käyntihaastatteluin, ja vuodesta 2021 alkaen työvoimatutkimuksen vastaajilla on ollut mahdollisuus vastata myös verkossa. Uusi EU-lainsäädäntö vaikutti otoksen muodostamiseen ja tulosten laskentatapaan eli estimointimenetelmään, sillä tutkimus kattaa nykyisin 15–89-vuotiaat, kun aiemmin kohderyhmänä oli 15–74-vuotiaat. Lisäksi estimointimenetelmässä huomioitiin tulosten luotettavuudelle asetetut laatuvaatimukset. Estimointimenetelmää uudistettaessa pyrittiin huomioimaan vastauskadon kasvu sekä sen vaikutus vastaajajoukon ikä- ja koulutusrakenteisiin.

Vastauskato voi vinouttaa vastaajajoukkoa, jos tutkimukseen vastaamattomuus ei ole satunnaista, vaan eri väestöryhmät vastaavat tutkimukseen eri todennäköisyydellä. Vaara tutkimustulosten vääristymiseen on olemassa, mikäli vastaamisalttiuteen liittyvät tekijät ovat yhteydessä myös tutkittavaan ilmiöön tai vastaushalukkuus riippuu suoraan tutkittavista asioista, esimerkiksi siitä, onko henkilö työllinen vai ei.

EU-lainsäädännön muutokset ja vastauskadon vaikutus huomioidaan muodostettaessa painokertoimia, joita käyttämällä vastanneiden antamat tiedot yleistetään koskemaan uuden määrittelyn mukaista perusjoukkoa. Painotuksessa hyödynnetään entistä laajemmin rekisterilähteistä saatavaa lisäinformaatiota, joka auttaa oikaisemaan vastauskadon aiheuttamaa vinoumaa ja kaventamaan estimaatin virhemarginaalia. Rekisteritietoa voi tarvittaessa käyttää myös vertailutietona. Estimointi paranee sitä enemmän, mitä suurempi on tiedonkeruussa kerätyn tiedon ja samaa ilmiötä kuvaavan rekisteritiedon välinen riippuvuus.

Uudistetussa painotuksessa käytetään perusjoukon väestörekisteritietojen (sukupuoli, ikä, alue ja kieli) ohella lisäinformaatiota seuraavista rekisterilähteistä: työ- ja elinkeinoministeriön työnvälitystilasto (rekisterityöttömyys), Tilastokeskuksen tutkintorekisteri (koulutusaste) ja työssäkäyntitilasto (pääasiallinen toiminta) sekä Verohallinnon tulorekisteri (kuukausittainen palkkasumma).

Aikasarjojen taannehtivat korjaukset

Menetelmä- ja määritelmämuutokset vaikuttivat useimpiin työvoimatutkimuksen estimaatteihin. Tulosten aikavertailun vuoksi myös aiempien vuosien luvut eli aikasarjat on tuotettu korjattuina vuodesta 2009 alkaen. Vertailukelpoiset aikasarjat on muodostettu takautuvasti käyttämällä uudella laskentatavalla tuotettua painokerrointa.

Myös uudistetusta EU-lainsäädännöstä johtuvat määritelmämuutokset oli pyritty ottamaan huomioon aikasarjakorjauksissa. Työllisen määritelmän merkittävimmät muutokset koskevat vanhempainvapaalla olevia sekä yrittäjiä ja yrittäjäperheenjäseniä.

Vuodesta 2021 alkaen ansiosidonnaisella vanhempainvapaalla olevat luokitellaan aina työllisiksi, mikä vaikuttaa etenkin äitien työllisyyteen. Aiemmin vanhempainvapaalla olevia ei laskettu työllisiksi, jos poissaolo työstä oli kestänyt yli 3 kuukautta. Määritelmämuutosten osalta aikasarjakorjauksissa hyödynnettiin vuosien 2009–2020 lomakkeilta johdettuja tietoja. Muutos kasvatti työllisten määrän estimaattia kyseisinä vuosina noin 12 000–15 000 henkilöllä.

Myös työllisen määritelmän muutos, joka koskee yrittäjiä, oli otettu huomioon vuosien 2009–2020 aikasarjojen korjauksessa. Vuodesta 2021 alkaen työstä poissa olleet yrittäjät luokitellaan työllisiksi tai ei-työllisiksi samoilla kriteereillä kuin palkansaajat. Aiemmin kaikki yrittäjät oli luokiteltu työllisiksi poissaolon pituudesta ja syystä riippumatta. Muutos pienensi työllisten määrän estimaattia noin 2 000–3 000 henkilöllä vuositasolla.

Uudistuksen vaikutukset eri aikasarjoihin

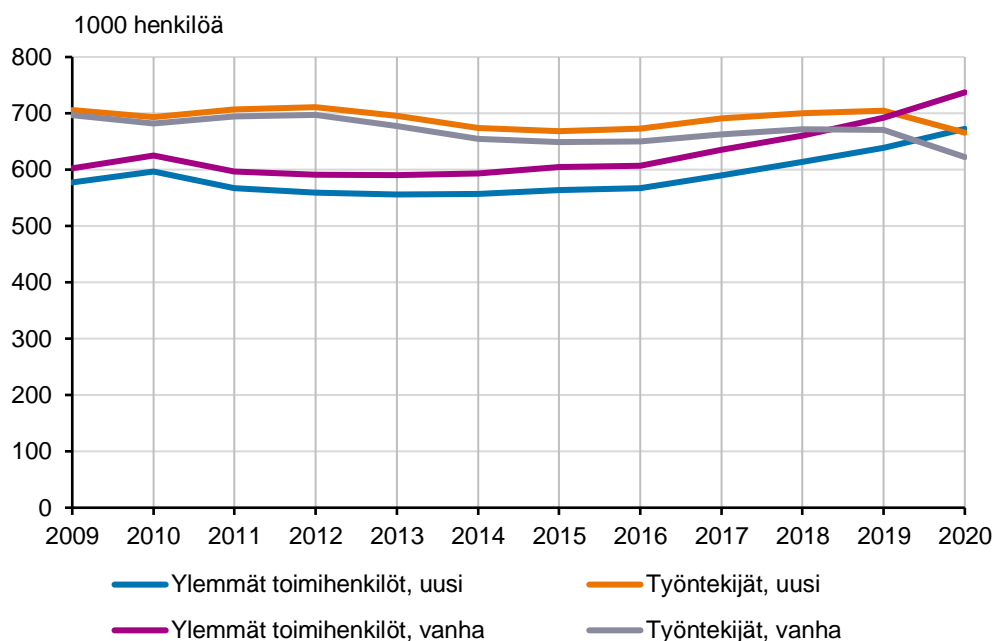
Työmarkkina-asema

Vuosille 2009–2020 tehdyn aikasarjakorjauksen aiheuttamat muutokset työttömien estimaatteihin olivat suhteellisesti vähäisiä, useimpina vuosina alle prosentin. Työvoiman ulkopuolella olevien määrät kasvoivat 2–3 %.

Työllisten aikasarjoihin uudistus aiheutti yleisesti pienehköjä, noin 1–2 % suuruisia muutoksia alaspäin. Työllisten määrän vähentyminen on tapahtunut pääasiassa palkansaajien ryhmässä.

Työllisten sosioekonomisen aseman mukaisissa alasarjoissa selvin muutos on ylempien toimihenkilöiden aikasarjassa, jossa uudet estimaatit ovat vanhan menetelmän mukaisia pienemmät. Enimmillään muutos on -9 % vuonna 2020. Työntekijöiden määrät puolestaan ovat kasvaneet muutaman prosentin verrattuna vanhan menetelmän mukaiseen aikasarjaan. Tätä selittää vastaajien koulutustason huomioiminen painojen kalibroinnissa.

Ylemmät toimihenkilöt ja työntekijät vuosina 2009–2020 vanhalla ja uudella menetelmällä laskettuna



Yrittäjien estimaateissa ei tapahtunut merkittäviä muutoksia.

Työajat

Ylempien toimihenkilöiden ja työntekijöiden määrien muutokset heijastuvat myös palkansaajien työaikojen pituutta kuvaaviin aikasarjoihin. Uudistuksen vaikutukset työllisten työaikoja käsitteleviin aikasarjoihin ovat olleet muutoin hyvin vähäisiä.

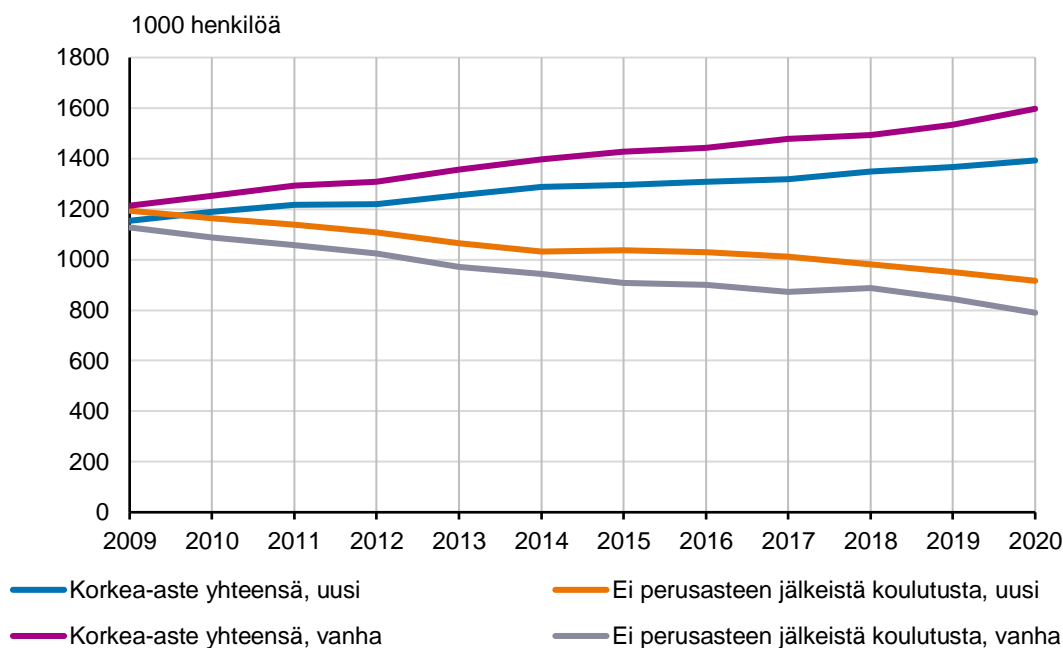
Tehtyjä vuosityöaikoja tarkastellessa eroja vanhojen ja uusien aikasarjojen välillä esiintyy erityisesti julkisen sektorin luvuissa. Kunta-alalla uuden aikasarjan työllisten ja palkansaajien tekemät työtunnit olivat noin 2–4 % pienemmät vanhaan sarjaan verrattuna ja valtiolla noin 2–7 % pienemmät. Yksityisellä sektorilla suuria muutoksia ei havaittu.

Koulutus

Vuosien varrella vastauskadon kasvu erityisesti vähemmän kouluttautuneiden joukossa on johtanut korkeakoulutettujen yliedustukseen vastausaineistossa. Uudessa painotusmenetelmässä hyödynnetään aputietona tutkintorekisteriä ja saadaan näin oikaistua koulutusasteen mukaista vastausaineiston vinoumaa aiempaa paremmin. Heikommin vastaavia ryhmiä edustavat vastaajat saavat aineistossa suuremman painon, joten aineistosta estimoitu väestön koulutusjakauma vastaa aiempaa paremmin todellisuutta ja on linjassa rekisteritietojen kanssa.

Uudistuksella oli selvä vaikutus tuloksiin koulutusasteen mukaan. Uuden menetelmän mukaisissa aikasarjoissa korkeakouluttautuneita oli työikäisessä väestössä 205 000 vähemmän (-13 %) vuonna 2020 verrattuna vanhan menetelmän mukaiseen estimaattiin. Heitä, joilla ei ole tiedossa perusasteen jälkeistä koulutusta, on vastaavasti 126 000 enemmän (+16 %) ja toisen asteen koulutuksen suorittaneita 78 000 enemmän (+4 %).

Korkea-asteen koulutuksen suorittaneet ja ilman perusasteen jälkeistä tutkintoa olevat 15–74-vuotiaat vuosina 2009–2020 vanhalla ja uudella menetelmällä laskettuna



On myös hyvä huomata, että vuoden 2021 alusta työvoimatutkimuksen koulutustiedoissa alettiin käyttää enemmän tutkintorekisteritietojen lisäksi lomakkeelta saatuja vastauksia, mikä vaikuttaa aikasarjojen vertailukelpoisuuteen. Vastaustietojen hyödyntäminen täydentää erityisesti ulkomailta muuttaneen väestön koulutustietoja.

Erot rekisteriin perustuvassa ja vastaustiedoilla täydennetyssä koulutusasteen jakaumissa ovat kohtalaisen pieniä. Osa vailla tutkintotietoa olevista siirtyy luonnollisesti vastaustiedon perusteella toisiin luokkiin.

Kansalaisuus

Vuoden 2021 alusta työvoimatutkimuksen painojen kalibroinnissa käytetään koulutusasteen lisäksi uutena aputietona Tilastokeskuksen väestötietokannasta saatavaa tietoa henkilön äidinkielestä. Tämän tiedon käyttö painokertoimien laskennassa auttaa korjaamaan muita kuin kotimaisia kieliä puhuvan väestön heikkoa vastausosuutta ja sen aiheuttamaa vastausaineiston vinoumaa. Näin aineistosta estimoidut ulkomaalaistaustaisen väestön jakaumat vastaavat paremmin todellisuutta.

Tämän uudistuksen vaikutus näkyy kansalaisuuden mukaisissa aikasarjoissa. Otoksen koko huomioiden, tietoja on mahdollista julkaista jaottelulla Suomen kansalaiset ja muut.

Piilotyöttömyys

Taannehtivista korjauksista huolimatta osassa aikasarjoista esiintyy katkoksia vuoden 2021 uudistuksen myötä. Yksi näistä on piilotyöttömyys, jota oli käytetty täydentämään kansainvälisen työjärjestö ILO:n määritelmää työttömyydestä. Eurostat on kehittänyt uusia työvoiman alikäyttöä kuvaavia indikaattoreita, jotka kuvaavat ilmiötä pelkkien piilotyöttömien tilastointia laajemmin. Näistä syistä Tilastokeskuksen työvoimatutkimuksesta ei enää julkaista erikseen tietoja piilotyöttömyydestä. Eurostatin uusiin indikaattoreihin voi tutustua [Eurostatin kotisivuilla \(englanniksi\)](#).

Vastaustavan merkitys

Tuloksiin voi vaikuttaa myös vastaustapa eli se, haastatellaanko henkilöä puhelimitse tai paikan päällä vai vastaako hän itsenäisesti verkossa. Erityisesti verkkovastauksiin voivat vaikuttaa haastattelua herkemmin tietotekniset ongelmat, henkilöiden vastausväsymys sekä kysymysten erilaiset tulkinnat. Näiden haasteiden vaikutus voidaan työvoimatutkimuksen aikasarjoissa havaita siinä, että tuntemattomien ryhmät ja 'en osaa sanoa' -vastaukset ovat lisääntyneet verkkolomakkeen käyttöönoton jälkeen, ja verkkovastaajat antavat haastateltuja useammin sisällöltään ristiriitaisia vastauksia.

Tiettyjen kysymysten vastausjakaumat voivat muutenkin olla erilaisia riippuen henkilöiden vastaustavasta. Tällöin on kuitenkin tärkeä tutkia, että johtuvatko eroavaisuudet annetuissa vastauksissa itse vastaustavasta vai paremminkin verkossa ja haastattelussa vastanneiden erilaisista taustatekijöistä, kuten iästä, sukupuolesta ja koulutustasosta.

Eurostatin julkaisemat aikasarjat

Eurostatille on toimitettu uudistetun työvoimatutkimuksen mukaiset takautuvasti korjatut neljännesvuosiaikasarjat 2009–2020 perusindikaattoreista, joihin kuuluvat estimaatit työmarkkina-asemasta sukupuolen ja iän mukaan. Lisäksi toimitettiin korjattuja aikasarjoja työmarkkina-asemasta sukupuolen ja koulutustason mukaan. Eurostat on julkaissut nämä aikasarjat omissa tietokannoissaan.

Edellä mainittujen tietojen lisäksi Eurostat on julkaissut muita aikasarjoja, joihin se on estimoinut takautuvat korjaukset omilla menetelmillään. Niissä voi siksi olla eroja Suomessa julkaistuihin estimaatteihin verrattuna. Eurostatin itse korjaamiin aikasarjoihin sisältyy tietoja muun muassa palkansaajista sekä työn ja koulutuksen ulkopuolella olevista nuorista.