

Laaturaportti: Työvoimatutkimus, kotitalousosa

Tilaston kuvaus

Työvoimatutkimus on otostutkimus, jonka avulla tilastoidaan 15–89-vuotiaan väestön työmarkkinoille osallistumista, työllisyyttä, työttömyyttä ja työaikaa kuukausittain, neljännesvuosittain ja vuosittain. Tutkimuksen tietosisältö perustuu EU:n asetukseen, ja tutkimuksen otokseen kuuluu joka kuukausi noin 12 500 henkilöä.

Työvoimatutkimus antaa tuoreen ja kattavan kuvan työvoimasta ja työmarkkinoiden muutoksista. Julkisuudessa seurataan kuukausittain erityisesti työllisyyden ja työttömyyden muutoksia edellisen vuoden vastaavasta kuukaudesta. Kausitasoitettuja aikasarjoja ja trendiaikasarjoja käytetään seurattaessa pitkän aikavälin kehitystä ja suhdannevaihteluita. Tutkimus antaa tietoa myös työikäisistä henkilöistä, jotka eivät ole työssä tai työnhakijoina.

Vuoden 2021 alussa työvoimatutkimuksen tietosisältöä, tiedonkeruuta ja estimointimenetelmää uudistettiin. Kyselylomakkeen sisältöä yhdenmukaistettiin aikaisempaa tarkemmin eri EU-maiden välillä tietojen vertailtavuuden parantamiseksi. Tietosisältö myös laajeni, kun kyselyyn lisätiin joitakin uusia kysymyksiä. Lisäksi uutena ikäryhmänä tutkimukseen otettiin mukaan 75–89-vuotiaat. Tämän ikäryhmän otoskoko on pieni ja sen vuoksi ikäryhmästä ei julkaista tietoja kuukausi- ja vuosineljänneksitasolla. Työvoimatutkimuksen kuukausi- ja neljännesvuositiedot koskevat edelleen 15–74-vuotiaiden ikäryhmää.

Tutkimuksen tiedonkeruutapaa uudistettiin tarjoamalla vastaajille mahdollisuus vastata puhelinhaastattelujen lisäksi myös verkkolomakkeella. Otoksen muodostamisessa ja tulosten laskentatavassa huomioitiin uusi EU-lainsäädäntö, vastauskadon lisääntyminen ja tiedonkeruutapaan toteutetut muutokset.

Vastaajien antamien tietojen pohjalta luodaan kuva koko 15–89-vuotiaan väestön toiminnasta työmarkkinoilla yhden viikon aikana. Työmarkkina-aseman perusluokittelussa väestö jaetaan työllisiin, työttömiin ja työvoiman ulkopuolisiin. Näistä työlliset ja työttömät yhdessä muodostavat työvoiman.

Tutkimus on paneelitutkimus, jossa samaa henkilöä haastatellaan viisi kertaa. Haastattelut tehdään kolmen kuukauden välein, paitsi neljäs haastattelu, joka tehdään kuuden kuukauden kuluttua kolmannesta haastattelusta. Ensimmäisen ja viimeisen haastattelun väli on 15 kuukautta. Tutkimuksen tietosisältö vaihtelee eri tutkimuskerroilla. Ensimmäisellä tutkimuskerralla kysytään tarkemmin joitakin perustietoja, joiden ennallaan pysyminen tarkistetaan seuraavilla tutkimuskerroilla. Viimeisellä eli viidennellä tutkimuskerralla tietosisältö on laajempi, ja perustutkimuksen lisäksi kyselyyn kuuluu myös kotitalousosa ja osio, jonka aihe vaihtuu vuosittain.

Tässä dokumentissa keskitytään työvoimatutkimuksen kotitalousosan kuvaukseen. Tarkemmat tiedot työvoimatutkimuksesta löytyvät tilaston [dokumentaationsivulta](#).

Yleiskuvaus

Työvoimatutkimuksen kotitalousosan tietoja on kerätty vuodesta 2003. Kotitaloushaastattelun aikana selvitetään kohdehenkilön kotitalouden jäsenet sekä kotitalouteen kuuluvien 15–89-vuotiaiden toiminta työmarkkinoilla. Jotta aineisto kuvaisi koko väestöä, siihen on lisätty ylimääräisen otoksen avulla myös sellaisia kotitalouksia, joiden kaikki jäsenet ovat 90 vuotta täyttäneitä. Haastattelutietojen lisäksi tietoja saadaan Tilastokeskuksen käytössä olevista hallinnollisista rekistereistä, mm. tutkintorekisteristä ja verorekisteristä.

Alle 15-vuotiaita lapsia ja yli 89-vuotiaita ei työvoimatutkimuksessa haastatella, vaan heille luodaan muutamia työmarkkina-asemaa kuvaavia tietoja rekisteritietojen perustella.

Kotitalousaineiston avulla voidaan tarkastella esimerkiksi lasten iän tai lukumäärän yhteyttä vanhempien työmarkkina-asemaan tai puolisoiden työaikajärjestelyitä.

Luokitukset

Työvoimatutkimuksessa käytettäviä luokituksia ovat Toimialaluokitus 2008 (TOL 2008, NACE Rev. 2), Ammattiluokitus 2010 (ISCO-08), Sosioekonomisen aseman luokitus 1989 (Ammattiluokitus 2010 pohjalta) ja Koulutusluokitus 2016 (ISCED 2011) sekä alueluokituksista suuralue, aluehallintovirasto (AVI), elinkeino-, liikenne- ja ympäristökeskus (ELY-keskus) sekä maakunta. Tilastovuodesta 2020 alkaen työvoimatutkimuksessa on otettu soveltuvin osin käyttöön Sektoriluokitus 2012 - luokitusstandardi.

Kattavuus

Työvoimatutkimuksen kotitalousosan vastaajien antamien tietojen pohjalta luodaan kuva koko Suomen 15–89-vuotiaan väestön toiminnasta työmarkkinoilla vuositasona. Kotitalousosan aineisto ei sisällä asevelvollisia.

Tilastolliset käsitteet ja määritelmät

Tutkimuksessa käytetyt käsitteet ja määritelmät noudattavat YK:n kansainvälisen työjärjestön ILO:n suosituksia ja Euroopan unionin tilastotoimen asetuksia.

Työllinen on henkilö, joka on tutkimusviikolla tehnyt työtä vähintään tunnin palkkaa tai yrittäjätuloa saadakseen. Myös tutkimusviikolla työstä tilapäisesti pois ollut henkilö lasketaan työlliseksi, jos poissaolon syy on äitiys- tai isyyysvapaa, ansiosidonnainen vanhempainvapaa, oma sairaus, loma tai työaikajärjestely; tai jos poissaolo kestää alle kolme kuukautta.

Työllisen määritelmän keskeisimmät muutokset vuoden 2021 alusta:

Vanhempainvapaalla olevat, jotka saavat ansiosidonnaista tukea, luokitellaan työllisiin, vaikka työstä poissaolo olisi kestänyt yli 3 kuukautta.

Yrittäjä tai samassa kotitaloudessa asuvan perheenjäsenen yrityksessä palkatta työskentelevä lasketaan työlliseksi vain, jos työstä poissaolo on kestänyt alle kolme kuukautta. Aiemmin yrittäjät laskettiin työlliseksi riippumatta poissaolon pituudesta tai syystä.

Työtön on henkilö, joka tutkimusviikolla on työtä vailla (ei ollut palkkatyössä tai tehnyt työtä yrittäjänä), on etsinyt työtä aktiivisesti viimeisen neljän viikon aikana palkansaajana tai yrittäjänä ja voisi aloittaa työn kahden viikon kuluessa. Myös henkilö, joka on työtä vailla ja odottaa sovitun työn alkamista kolmen kuukauden kuluessa, luetaan työttömäksi, jos hän voisi aloittaa työn kahden viikon kuluessa. Työttömäksi luetaan myös työpaikastaan toistaiseksi lomautettu, joka täyttää edellä mainitut kriteerit.

Työvoimaan kuuluvat kaikki ne henkilöt, jotka tutkimusviikolla olivat työllisiä tai työttömiä.

Työvoiman ulkopuolella olevilla tarkoitetaan henkilöitä, jotka tutkimusviikolla eivät olleet työllisiä tai työttömiä. Työvoiman ulkopuolella olevista voidaan käyttää myös nimitystä työvoimaan kuulumattomat.

Lapsiperhe on perhe, johon kuuluu vähintään yksi kotona asuva alle 18-vuotias lapsi. Lapsiperheen vanhempia - äitejä ja isiä - ovat ne naiset ja miehet, joilla on samassa kotitaloudessa asuvia alle 18-vuotiaita omia tai puolison lapsia.

Lapsettomiksi luokitellaan tässä naiset ja miehet, joilla ei ole samassa kotitaloudessa asuvia alle 18-vuotiaita lapsia. Lapsettomiin kuuluvat siis

- henkilöt, joilla ei ole lainkaan lapsia
- sekä vanhemmat, joiden lapset ovat täyttäneet 18 vuotta tai ovat jo muuttaneet omaan kotitalouteen.

Muita työvoimatutkimuksen käsitteitä on selitetty [dokumentaationsivulla](#).

Tilastoyksikkö

Työvoimatutkimuksen tilastoyksikkö on henkilö.

Tilaston perusjoukko

Työvoimatutkimuksen perusjoukon muodostavat Suomessa vakinaisesti asuvat 15–89-vuotiaat henkilöt. Perusjoukkoon kuuluvat myös tilapäisesti ulkomailla (alle vuoden) oleskelevat sekä Suomen väestötietojärjestelmään rekisteröidyt ulkomaalaiset, joiden oleskelu Suomessa kestää vähintään vuoden.

Työvoimatutkimuksen otos poimitaan ositetulla systemaattisella otannalla väestön keskusrekisteriin perustuvasta Tilastokeskuksen väestötietokannasta kahdesti vuodessa.

Kotitalousotos muodostetaan työvoimatutkimuksen otokseen poimituille henkilöille ennen 5. tutkimuskertaa. Väestökehikon asuinpaikkatunnuksen avulla työvoimatutkimuksen 15–89-vuotiaiden muodostetaan asuntokunnat, joista tarkistusvaiheen jälkeen poistetaan kotitalouteen kuulumattomat henkilöt. Kotitalousosan otos muodostetaan neljännesvuosittain.

EU-asetuksen mukaan vuodesta 2021 alkaen hyväksytyksi kotitaloudeksi määritellään ne taloudet, joissa kaikki jäsenet ovat vastanneita. Aiemmin hyväksytyksi kotitaloudeksi on määritelty ne taloudet, joissa enintään puolet jäsenistä on kadossa.

Viitealue

Työvoimatutkimuksen viitealue on koko Suomi.

Ajallinen kattavuus

Työvoimatutkimus uudistettiin vuoden 2021 alussa, jolloin kotitalousosan aikasarjat vuosilta 2013–2020 korjattiin takautuvasti uudistetun estimointimenetelmän mukaisiksi. Korjausmenettelyssä työllisen määritelmämuutos oli huomioitu 15–74-vuotiaiden naisten koko kohderyhmän tasolla. Sen sijaan naisten osajoukkotasolla aikasarjakorjattuihin työllisyystietoihin liittyy korjausmenettelyssä johtuen epävarmuutta, eivätkä aiempien

vuosien tiedot ole täysin vertailukelpoisia vuodesta 2021 lähtien tuotettujen lukujen kanssa.

Mittayksikkö

Työvoimatutkimuksen mittayksiköjä ovat henkilöiden määrä (1 000 henkilöä) ja työtuntien määrä (1 000 tuntia).

Viiteajankohta

Pääosa työvoimatutkimuksen kysymyksistä koskee tutkimuskuukauden yhtä viikkoa. Tiedot kootaan kysymällä kaikilta haastateltavilta heidän toimintaansa ko. viikolla. Työvoimatutkimuksen tiedonkeruussa käytetään jatkuvaa tutkimusviikkoa. Tämä tarkoittaa, että vuoden kaikki viikot ovat edustettuina tutkimuksessa ja otos jaetaan tasaisesti kaikille viikoille. Vuositulokset kuvaavat tilannetta ko. tutkimusjakson ”keskimääräisellä” viikolla.

Toimintavaltuudet

Lainsäädäntö ja muut sopimukset

Tilastojen laadintaa ohjaa tilastolaki. Tilastolaissa säädetään muun muassa tiedonkeruusta, tietojen käsittelystä ja tiedonantovelvollisuudesta. Tietojen käsittelyyn tilastoja tuotettaessa sovelletaan tilastolain lisäksi tietosuojalakia sekä lakia viranomaisen toiminnan julkisuudesta. Tilastokeskus soveltaa tilastoja laatiessaan EU:n tilastosäädöksiä, jotka ohjaavat kaikkien EU-maiden tilastovirastoja.

Lisätietoja: [Tilastolainsäädäntö](#).

Tietojen jakaminen

Työvoimatutkimuksen vuositiedot, mukaan lukien kotitalousosan tiedot, toimitetaan Eurostatin noin kolmen kuukauden viiveellä tutkimusvuoden päättymisestä.

Tilastollinen tietosuoja

Tietosuojaperiaatteet

Tilastotarkoituksiin kerätyn tiedon tietosuoja taataan tilastolain (280/2004), viranomaisten toiminnan julkisuudesta (621/1999) annetun lain, EU:n yleisen tietosuoja-asetuksen (EU) 2016/679 ja tietosuojalain (1050/2018) vaatimusten mukaisesti.

Tietoaineistot on suojattu käsittelyn kaikissa vaiheissa tarvittavin fyysisin ja teknisin ratkaisuin. Tilastokeskus on laatinut yksityiskohtaiset määräykset ja ohjeet tietojen luottamukselliseen käsittelyyn. Henkilökunnalla on pääsy vain työtehtävien kannalta välttämättömiin tietoihin. Tiloihin, joissa yksikötason aineistoa käsitellään, ei ulkopuolisilla ole pääsyä. Henkilökunnan jäsenet ovat allekirjoittaneet salassapitositoumuksen palvelukseen tullessaan. Tietosuojaan rikkomisesta seuraa rangaistus.

Lisätietoja: [Tietosuoja Tilastokeskuksessa](#).

Tietosuoja ja -turva tietoja käsiteltäessä

Työvoimatutkimuksen tietoja ei luovuteta tunnistettavassa muodossa Tilastokeskuksen ulkopuolelle (Tilastolaki 280/2004, Henkilötietolaki 523/1999). Aineiston luovuttaminen on mahdollista ainoastaan erillisen käyttölupapäätöksen perusteella ja tunnistetiedot poistettuna tieteellistä tutkimusta ja tilastollisia selvityksiä varten. EU-työvoimatutkimusta varten muokatut aineistot toimitetaan EU:n tilastovirastolle ilman tunnistetietoja. Työvoimatutkimuksen tiedot suojataan Tilastokeskuksen tietosuojaohjeessa määritellyn suojausluokan mukaan.

[Työvoimatutkimuksen tietosuojaseloste](#) on nähtävissä tietosuojaselosteiden sivulla.

Julkistamispolitiikka

Julkistamiskalenteri

Tilastokeskus julkistaa uutta tilastotietoa arkipäivisin kello 8:00 verkkopalvelussaan. Tilastojen julkaisuajankohdat kerrotaan ennakoon verkkopalvelusta löytyvässä julkistamiskalenterissa. Tiedot ovat julkisia sen jälkeen, kun ne ovat päivittyneet verkkopalveluun.

Lisätietoja: [Tilastojen julkistamisperiaatteet Tilastokeskuksessa](#).

Pääsy julkistamiskalenteriin

Tilastokeskuksen julkistamiskalenteri: [Tulevat julkaisut](#).

Tilaston tulevat julkaisut ovat tilaston sivuilla: [Tilaston tulevat julkaisut](#).

Käyttäjien käyttöoikeudet

Tiedot julkistetaan kaikille käyttäjille samanaikaisesti. Tilaston tietoja saa Tilastokeskuksessa käsitellä ja niistä saa antaa tietoja ennen julkistamista vain henkilö,

joka osallistuu kyseisen tilaston laadintaan tai tarvitsee ko. tilaston tietoja omassa työssään ennen tietojen julkistamista.

Lisätietoja: [Tilastojen julkistamisperiaatteet](#).

Tilastokeskus on aineistojen tuottaja ja tekijänoikeuden haltija, ellei tuotteen, tiedon tai palvelun yhteydessä erikseen toisin ilmoiteta. Tilastotietojen [käyttöehdot](#).

Jakelutiheys

Kotitalousosan vuositulokset julkistetaan määräpäivinä Internetissä [työvoimatutkimuksen kotisivulla](#). Julkaistavat tiedot ovat lopullisia.

Saatavuus ja selkeys

Kotitalousosan tilastotiedot julkaistaan tietokantataulukoina StatFin-tietokannassa. Tietokanta on tietojen ensisijainen julkaisupaikka, ja uudet tiedot päivitetään ensimmäisenä tietokantaan. Tilastotietojen julkistuksessa voidaan päivittää olemassa olevia tietokantataulukoita uusilla tiedoilla tai julkaista kokonaan uusia tietokantataulukoita.

StatFin-tietokannassa julkaistavien tilastotietojen rinnalla verkkopalvelussa julkaistaan yleensä tiedote keskeisimmistä tiedoista. Tiedotteet ja tietokantataulukot julkaistaan kolmella kielellä, suomeksi, ruotsiksi ja englanniksi. Tiedotteiden kieliversiot voivat olla suomenkielistä suppeampia.

Tietokantataulukoiden lisäksi keskeisimmät kotitalousosan tulokset julkaistaan määräpäivinä katsauksena verkkopalvelussa [työvoimatutkimuksen kotisivulla](#). Julkistuksiin ja tietokantataulukoihin liittyvistä aikataulumuutoksista ja korjauksista viestitään muutostiedotteilla verkkopalvelussa.

Julkaisut (laajemmat/muut)

Perheet ja työ -tiedote julkaistaan vuosittain Statfin-tietokantataulukoiden päivityksen yhteydessä [tilaston sivulla](#).

Tietokantataulukoiden lisäksi keskeisimmät kotitalousosan tulokset julkaistaan määräpäivinä katsauksena verkkopalvelussa [työvoimatutkimuksen kotisivulla](#). Julkistuksiin ja tietokantataulukoihin liittyvistä aikataulumuutoksista ja korjauksista viestitään muutostiedotteilla verkkopalvelussa.

Verkkotietokanta (StatFin-tietokanta)

Kotitalousosan tietokantataulukot 13nx ja 13ny ovat saatavissa [työvoimatutkimuksen StatFin-tietokannasta](#).

Yksikkötason aineistojen saatavuus

Työvoimatutkimuksen tietoja ei luovuteta tunnistettavassa muodossa Tilastokeskuksen ulkopuolelle (Tilastolaki 280/2004, Henkilötietolaki (523/1999)). Vastausaineistoa käytetään ainoastaan tilastollisiin tarkoituksiin. Aineiston luovuttaminen on mahdollista ainoastaan erillisen käyttö lupapäätöksen perusteella ja tunnistetiedot poistettuna tieteellistä tutkimusta ja tilastollisia selvityksiä varten.

Menetelmädokumentointi

Työvoimatutkimuksen menetelmäselosteet löytyvät tilaston [dokumentaationsivuilta](#).

Laadunhallinta

Laadunvarmistus

Laadunhallinta edellyttää toiminnan kokonaisvaltaista ohjausta. Tilastoalan oma laadunhallinnan kehikko on Euroopan tilastojen käytännestännöstö (CoP). Kehikot täydentävät toisiaan. Myös Suomen virallisen tilaston laatukriteerit ovat yhteensopivat Euroopan tilastojen käytännestäntöjen kanssa. Lisätietoja: [Laadunhallinta](#). Tilastokeskuksessa tehdään vuosittain tilastojen läpivalaisuja, joilla osaltaan varmistetaan tilastojen laatua. Työvoimatutkimuksen läpivalaisu toteutettiin lokamarraskuussa 2011 ja tilastotuotannon laadunvarmistus lokakuussa 2016.

Laadun arviointi

Työvoimatutkimuksen laatua arvioidaan useassa tilastoprosessin eri vaiheessa.

Relevanssi

Käyttäjien tarpeet

Tutkimuksen tuloksia käytetään muun muassa työvoimapolitiittisten ennusteiden ja suunnitelmien laadinnassa, päätösten teon tukena sekä erilaisten toimien työllisyysvaikutusten seurannassa. Tärkeitä tulosten käyttäjiä ovat ministeriöt, työnantaja- ja työntekijäjärjestöt, yliopistot ja tutkimuslaitokset, kansainväliset järjestöt ja Euroopan unioni. Euroopan unionin tilastovirasto, Eurostat, ohjaa tutkimuksen sisältöä ja seuraa sen laatua.

Tarkkuus ja luotettavuus

Tarkkuus ja luotettavuus yleisesti

Työvoimatutkimuksen lukujen luotettavuuteen vaikuttavat mittausvirheet, otannasta aiheutuva vastauskato sekä satunnaisvaihtelu. Mittausvirheitä voi syntyä siitä, että kysymyksiä voidaan ymmärtää ja tulkita eri tavoin tai että vastaajat jättävät kertomatta jotain. Mittausvirheitä pyritään vähentämään haastattelijakoulutuksen sekä lomakkeen testauksen avulla. Vastauskato voidaan jakaa yksikkö- ja erä- eli osittaiskatoon. Yksikkökato tarkoittaa, että otoshenkilöltä ei saada lainkaan haastattelua esimerkiksi kieltäytymisen tai tavoittamattomuuden takia. Yksikkökatoa voidaan korjata painokertoimien avulla. Eräkadolla taas tarkoitetaan kysymyskohtaista katoa. Tällöin haastattelu on saatu, mutta joissakin kysymyksissä esiintyy puutteita esimerkiksi haastattelun keskeytymisen tai vastaushaluttomuuden vuoksi.

Otannasta aiheutuvalla satunnaisvaihtelulla tarkoitetaan sitä, että eri otoksista lasketut luvut poikkeaisivat jonkin verrantoisistaan. Arvioitaessa karkeasti otannasta aiheutuvan satunnaisvaihtelun suuruutta erilaisissa tilanteissa on pääperiaatteena, että lukuihin sisältyy sitä vähemmän otannasta aiheutuvaa epävarmuutta, 1) mitä suuremman otoksenpohjalta luvut on laskettu ja 2) mitä suurempaa väestöryhmää luvut kuvaavat. Esimerkiksi neljännesvuotta koskevat luvut ovat tarkempia kuin samaa asiaa koskevat kuukausiluvut, koska neljännesvuositiedot on saatu haastattelemalla kolminkertainen määrä henkilöitä verrattuna kuukausitietoihin. Koko vuotta koskevat luvut ovat kaikkein tarkimpia. Jälkimmäinen periaate tarkoittaa sitä, että samankokoisella otoksella saadut työllisten ja työttömien lukumääriä koskevat luvut eli estimaatit ovat sitä tarkempia, mitä suurempaa osaryhmää ne koskevat. Tarkasteltavan osaryhmän pienentyessä otannasta aiheutuvan satunnaisvaihtelun osuus kasvaa. Tämän vuoksi esimerkiksi työttömien lukumäärät eri ikäryhmissä tai eri alueilla eivät ole yhtä luotettavia kuin kaikkien työttömien lukumäärä.

Työvoimatutkimuksen estimaattien laatua arvioidaan vuosittain ja julkaistaville estimaateille asetetaan rajat. Rajojen avulla estimaattien käytöstä varoitetaan tai ne peitetään julkaistavissa tiedoissa. Vuosiestimaattien julkistamisessa noudatamme Eurostatin ohjetta. Alle 3000 henkilön estimaatit peitetään taulukoissa automaattisesti pisteellä (.) ja alle 7000 henkilön estimaatista varoitetaan. Rajat koskevat pyöristämättömiä lukuja ja näiden lukujen pohjalta laskettuja prosenttiosuuksia.

Otantavirhe

Otannasta aiheutuvaa epätarkkuutta arvioidaan estimaatin **keskivirheen** avulla. Keskivirheen suuruuteen vaikuttavat sekä otoksen koko että tarkastelun kohteena olevan ominaisuuden vaihtelu eli muuttujan varianssi. Keskivirheen avulla voidaan laskea estimaatin **luottamusväli**, jolla haettu perusjoukon arvo sijaitsee tietyllä todennäköisyydellä. Työvoimatutkimuksessa käytetty **95 prosentin luottamusväli on väli, jolla kiinnostuksen kohteena olevan ominaisuuden todellinen arvo sijaitsee 95 prosentin todennäköisyydellä**. Jos esimerkiksi tietyn kuukauden työttömien määrän estimaatti on 230 000 ja sen keskivirhe 7 700, on työttömien määrän 95 prosentin luottamusväli $230\,000 \pm 15\,100$ eli 214 900–245 100 henkeä. Estimaattiin lisättävä ja siitä vähennettävä osa saadaan kertomalla estimaatin keskivirhe 95 prosentin luottamusvälin kertoimella 1,96. Tämä osa kuvaa otannasta aiheutuvaa epävarmuutta ja sitä kutsutaan estimaatin **virhemarginaaliksi**.

Esimerkkejä työllisten ja työttömien lukumäärän tarkkuudesta osaryhmän koon mukaan:

Satunnaisvaihtelun suuruusluokan havainnollistamiseksi esitetään seuraavissa taulukoissa 1–2 esimerkkejä erikokoisista työllisten ja työttömien lukumäärän estimaateista, niiden virhemarginaaleista ja muista luotettavuuden tunnusluvuista. Esimerkeissä esitetty satunnaisvaihtelun suuruusluokka on karkea arvio satunnaisvaihtelun ylärajasta, kun tarkastelun kohteena on **vastaavankokoinen työllisten tai työttömien lukumäärän estimaatti sukupuolen ja nuorimman lapsen iän mukaan**.

Taulukosta 1 voidaan esimerkiksi nähdä, että jos osaryhmän työllisten vuosiestimaatti on noin 56 000 henkeä, sijoittuu perusjoukon todellinen työllisten lukumäärä 95 prosentin todennäköisyydellä välille $56\,000 \pm 5\,200$ henkeä. Estimaatit ovat sitä tarkempia mitä suurempaa osaryhmää ne koskevat. Jos esimerkiksi verrataan kahta alle 18-vuotiaiden lasten äitien ryhmää ”nuorin lapsi alle 1-2-vuotias” ja ”nuorin lapsi 7-17-vuotias”, nähdään että ensimmäisen ryhmän estimaatin 95 prosentin luottamusväli, $56\,000 \pm 5\,200$ henkeä, on suhteessa estimaatin kokoon selvästi suurempi (suhteellinen keskivirhe 4,8 %) kuin jälkimmäisen ryhmän estimaatin kohdalla (suhteellinen keskivirhe 1,6 %).

Työllisten ja työttömien estimaattien koko ja luotettavuus vaikuttaa myös näiden estimaattien perusteella laskettaviin muihin tunnuslukuihin kuten työllisyys- ja työttömyysteiden estimaatteihin. Myös näiden ryhmien työllisyysasteiden suhteelliset keskivirheet eroavat huomattavasti toisistaan.

Taulukko 1. Esimerkkejä työvoimatutkimuksen kotitalousosan vuosiestimaattien tarkkuudesta: Alle 18-vuotiaiden lasten äidit, 20–59-vuotiaat

		Vuosiestimaatti	Vuosiestimaatin virhemarginaalit (95 %:n luottamusväli)	Keskivirhe	Suhteellinen keskivirhe, %
Työlliset, henkeä	Yhteensä	399 000	± 6 700	3 400	0,9
Nuorimman lapsen ikä	Alle 1-vuotias	24 000	± 3 600	1 900	7,6
	1–2-vuotias	54 000	± 5 500	2 700	4,8
	3–6-vuotias	83 000	± 6 200	3 200	3,8
	7–17-vuotias	238 000	± 7 500	3 900	1,6
Työttömät, henkeä	Yhteensä	20 000	± 3 400	1 700	9,0
Nuorimman lapsen ikä	Alle 3-vuotias	5 000	± 1 900	1 000	19,6
	3–6-vuotias	5 000	± 1 900	1 000	20,6
	7–17-vuotias	10 000	± 2 700	1 400	13,9
Työllisyysaste, %	Yhteensä	81,9	± 1,4	0,7	0,9
Nuorimman lapsen ikä	Alle 1-vuotias	61,3	± 7,0	3,6	5,8
	1–2-vuotias	67,8	± 4,8	2,5	3,6
	3–6-vuotias	81,2	± 3,3	1,7	2,1
	7–17-vuotias	89,4	± 1,5	0,8	0,9
Työttömyysaste, %	Yhteensä	4,6	± 0,8	0,4	8,8
Nuorimman lapsen ikä	Alle 3-vuotias	5,9	± 2,2	1,1	19,0
	3–6-vuotias	5,3	± 2,1	1,1	20,0
	7–17-vuotias	3,9	± 1,1	0,5	13,7

Taulukko 2. Esimerkkejä työvoimatutkimuksen kotitalousosan vuosiestimaattien tarkkuudesta:
Alle 18-vuotiaiden lasten isät, 20–59-vuotiaat

		Vuosiestimaatti	Vuosiestimaatin virhemarginaalit (95 %:n luottamusväli)	Keskivirhe	Suhteellinen keskivirhe (%)
Työlliset, henkeä	Yhteensä	397 000	± 5 000	2 600	0,6
Nuorimman lapsen ikä	alle 1-vuotias	34 000	± 4 600	2 300	6,9
	1–2-vuotias	69 000	± 6 400	3 200	4,7
	3–6-vuotias	84 000	± 6 300	3 200	3,8
	7–17-vuotias	210 000	± 7 600	3 900	1,8
Työttömät, henkeä	Yhteensä	13 000	± 2800	1 400	10,9
Nuorimman lapsen ikä	alle 3-vuotias	4 000	± 1800	900	22,7
	3–6-vuotias	2 300	± 1 500	700	30,5
	7–17-vuotias	6 600	± 2 300	1 100	16,2
Työllisyysaste, %	Yhteensä	90,7	± 1,1	0,6	0,6
Nuorimman lapsen ikä	alle 1-vuotias	87,5	± 5,1	2,6	3,0
	1–2-vuotias	90,4	± 3,2	1,6	1,8
	3–6-vuotias	90,9	± 2,8	1,4	1,6
	7–17-vuotias	91,2	± 1,7	0,8	0,9
Työttömyysaste, %	Yhteensä	3,1	± 0,7	0,3	10,7
Nuorimman lapsen ikä	alle 3-vuotias	3,7	± 1,6	0,8	22,3
	3–6-vuotias	2,7	± 1,6	0,8	30,0
	7–17-vuotias	3,1	± 1,0	0,5	16,1

Katovirhe

Kaikilta otokseen valituilta henkilöiltä ei saada vastausta, sillä joitakin henkilöitä ei koskaan tavoiteta ja monet kieltäytyvät vastaamasta tutkimukseen. Tätä kutsutaan vastauskadoksi. Vastauskato on yleistynyt viime vuosina. Estimointi perustuu vain vastanneisiin henkilöihin, joita on vastauskadon vuoksi vähemmän. Silloin virhemarginaali kasvaa ja luottamusvälitlevenevät. Vastauskato pyritään ottamaan huomioon estimoinnissa, esimerkiksi uudelleenpainotusmenetelmillä kalibrointiestimaattorin avulla.

Käsittelyvirhe

Tilastojulkistamisen mahdollisia virheitä ovat muun muassa julkistusten teksteissä, kuvioissa tai taulukoissa esitetyt virheelliset luvut tai väärin käsitteiden käyttö. Toisin kuin tietojen tarkentuminen, virhetilanteet ovat odottamattomia poikkeamia normaalista tilastotuotannosta.

Virheet korjataan ja niistä tiedotetaan käyttäjille mahdollisimman nopeasti. Tilastokeskus kertoo merkittävistä virheistä samassa laajuudessa ja samoja kanavia käyttäen kuin varsinaisten tietojen julkistamisessa.

Korjattuihin tilastojulkistuksiin lisätään merkintä korjauksesta sekä tieto korjauksen ajankohdasta. Mikäli mahdollista, myös alkuperäinen, virheellinen tieto jätetään näkyviin.

Tilastoissa esiintyvät kirjoitus- ja muut muotovirheet korjataan mahdollisimman nopeasti ja joustavasti, eikä niistä tehdä erillistä merkintää verkkosivuille.

Tilastotietokantoihin tehtävistä muutoksista uutisoidaan Muutoksia tietokannassa - sivulla. Tilastotietokannoissa ylläpidetään vain tuoreimpia tietoja.

Myös aikataulupoikkeamat lasketaan virhetilanteiksi. Mikäli tilastotiedon julkistaminen myöhästyy merkittävästi ilmoitetusta ajankohdasta, ilmoitetaan myöhästymisestä Tilastokeskuksen kotisivulla.

Oikea-aikaisuus ja täsmällisyys

Oikea-aikaisuus

Työvoimatutkimuksen kotitalousosan tulokset julkaistaan vuosittain noin viiden kuukauden kuluttua tutkimusvuoden päättymisestä.

Täsmällisyys

Julkaisukalenterin ja varsinaisen julkaisupäivän välillä ei esiinny viiveitä.

Johdonmukaisuus ja vertailukelpoisuus

Työvoimatutkimus uudistettiin vuoden 2021 alussa. Uuden estimointimenetelmän mukaiset aikasarjat alkavat vuodesta 2021. Työvoimatutkimus kotitalousosan aikasarjat vuosilta 2013–2020 korjattiin takautuvasti uuden estimointimenetelmän mukaisiksi. Korjausmenettelyssä työllisen määritelmämuutos oli huomioitu 15–74-vuotiaiden naisten koko kohderyhmän tasolla. Sen sijaan naisten osajoukkotasolla aikasarjakorjattuihin työllisyystietoihin liittyy korjausmenettelyssä johtuen epävarmuutta eivätkä ne ole täysin vertailukelpoisia vuodesta 2021 lähtien tuotettuihin lukuihin. Korjatut kotitalousosan

keskeiset estimaatit tuottavat yhdenmukaisia vuosiestimaatteja kuukausityövoimatutkimuksen kanssa.

Maantieteellinen vertailukelpoisuus

Työvoimatutkimuksen nykyinen tietosisältö perustuu pääosin Euroopan unionin työvoimatutkimuksia koskeviin asetuksiin (577/98, 2257/2003 ja 430/2005). Euroopan unionin työvoimatutkimuksesta löytyy tarkempi kuvaus osoitteesta:

[http://ec.europa.eu/eurostat/statistics-explained/index.php/Labour_market_and_Labour_force_survey_\(LFS\)_statistics](http://ec.europa.eu/eurostat/statistics-explained/index.php/Labour_market_and_Labour_force_survey_(LFS)_statistics).

Tutkimuksessa käytetyt käsitteet ja määritelmät noudattavat YK:n kansainvälisen työjärjestön ILO:n suosituksia ja Euroopan unionin tilastotoimen asetuksia, joten tilaston tiedot ovat kansainvälisesti vertailukelpoisia ja kattavat koko Suomen.

Ajallinen vertailukelpoisuus

Kuukausittainen työvoimatutkimus, aluksi nimellä työvoimatiedustelu, on ollut käynnissä vuodesta 1959 alkaen. Tänä aikana sen tietosisältöä, tiedonkeruutapoja ja menetelmiä on kehitetty useaan otteeseen. Keskeisimmistä tiedoista on vertailukelpoinen aikasarja vuodesta 1989 asti.

Suomen liittyttyä Euroopan unioniin työvoimatutkimus harmonisoitiin Euroopan unionin yhteiseen työvoimatutkimukseen. Aluksi vuosina 1995–1998 EU-työvoimatutkimuksen tiedot kerättiin erillisenä haastattelututkimuksena keväisin maaliskuu–toukokuussa. Kuukausitutkimus uudistettiin asteittain vastaamaan EU-työvoimatutkimusta. Vuonna 1997 kuukausitutkimuksen sisältö laajentui, tiedonkeruu muuttui tietokoneavusteiseksi puhelinhaastatteluksi (CATI) sekä käsitteet ja määritelmät yhdenmukaistettiin vastaamaan entistä paremmin Euroopan unionin ja YK:n työjärjestön ILO:n ohjeita ja suosituksia. Työttömän määritelmää tarkistettiin vielä toukokuussa 1998 ja julkaistut aikasarjat korjattiin vuodesta 1989 alkaen taannehtivasti uudistettujen määritelmien mukaisiksi.

Huhtikuussa 1999 työvoimatutkimuksen tietosisältö laajeni edelleen, kun kuukausitutkimus ja EU-työvoimatutkimusyhdistyivät yhdeksi jatkuvaksi työvoimatutkimukseksi. Vuodesta 1999 lähtien työvoimatutkimuksen yhteydessä tehtiin EU-maiden yhteinen ad hoc -lisätutkimus, jonka aihe vaihteli vuosittain.

Vuoden 2000 alusta siirryttiin jatkuvaan tietokeruuseen, jossa kaikki vuoden viikot ovat tutkimusviikkoja. Aiemmin kunkin kuukauden tiedot kerättiin yhdeltä tutkimusviikolta, yleensä kuukauden keskimmaiselta viikolta. Jatkuvaan tutkimusviikkoon siirtyminen vaikutti työpäivä- ja työtuntitietoihin, minkä vuoksi ne eivät ole vuodesta 2000 eteenpäin täysin vertailukelpoisia aiempiin tietoihin. Vaikutukset työllisten määrää kuvaaviin

lukuihin jäivät Tilastokeskuksessa tehdyn selvityksen mukaan pieniksi (Työvoimatilasto 2000).

Vuodesta 2003 alkaen tutkimuksen tietosisältö laajeni osaotokselta kerättävällä kotitalousosalla.

Vuonna 2008 otettiin käyttöön uusi tiedonkeruulomake: tutkimuksen tietosisältö muuttui hieman ja osa tiedoista siirryttiin keräämään osaotokselta, jolta kerättyjä tietoja käytetään vain vuositietoina. Myös joidenkin käsitteiden määritelmät muuttuivat. Uudistukset perustuivat EU:n asetuksiin, ja niillä pyrittiin parantamaan EU-maiden työvoimatilastojen vertailtavuutta.

Vuoden 2021 alussa työvoimatutkimuksen tietosisältöä, tiedonkeruuta ja estimointimenetelmää uudistettiin. Kyselylomakkeen sisältöä yhdenmukaistettiin aikaisempaa tarkemmin eri EU-maiden välillä tietojen vertailtavuuden parantamiseksi. Tietosisältö myös laajeni, kun kyselylomakkeelle lisätiin uusia kysymyksiä. Tärkeimmät muutokset koskevat kysymyksiä työajasta ja sitä, milloin henkilö määritellään työlliseksi. Jatkossa esimerkiksi vanhempainvapaalla olevat, jotka saavat ansiosidonnaista tukea, luokitellaan työllisiksi poissaolon pituudesta riippumatta. Lisäksi uutena ikäryhmänä tutkimukseen otettiin mukaan 75–89-vuotiaat. Tämän ikäryhmän otoskoko on pieni ja sen vuoksi ikäryhmästä ei julkaista tietoja kuukausi- ja vuosineljänneksellä. Työvoimatutkimuksen kuukausi- ja neljännesvuositiedot koskevat edelleen 15–74-vuotiaiden ikäryhmää. Tutkimuksen tiedonkeruutapaa uudistettiin tarjoamalla vastaajille mahdollisuus vastata puhelin- ja käyntihaastattelujen lisäksi myös vastausmahdollisuutta verkkolomakkeella. Kotitalousosan verkkotiedonkeruu aloitettiin vuotta myöhemmin eli vuonna 2022.

Euroopan unionin tilastoviraston, Eurostatin, julkaisemat Suomen työvoimatutkimuksen tulokset eroavat Suomessa julkaistuista siten, että asevelvolliset eivät ole lainkaan mukana Eurostatin tiedoissa. Useimmissa EU-maissa asevelvolliset eivät kuulu työvoimatutkimuksen kohdejoukkoon eli yksityisissä kotitalouksissa asuvaan väestöön. Tämä aiheuttaa eroja etenkin 15–24-vuotiaita koskeviin tuloksiin. Suomen julkaisemissa luvuissa asevelvollisuutta suorittavat ovat mukana työvoiman ulkopuolisessa väestössä. Kotitalousosan aineisto ei sisällä asevelvollisia vuodesta 2021 alkaen.

Otoksen muodostamisessa ja tulosten laskentatavassa huomioitiin uusi EU-lainsäädäntö, vastauskadon lisääntyminen ja tiedonkeruutapaan toteutetut muutokset. Kotitalousosan aikasarjat vuosilta 2013–2020 korjattiin taannehtivasti uuden estimointimenetelmän mukaisiksi.

Tietojen revisoituminen

Työvoimatutkimuksen kotitalousosan julkaistut luvut ovat lopullisia. Kertaalleen julkaistujen lukujen korjaaminen on ollut hyvin harvinaista.

Tietojen revisoitumislinjaukset

Jo julkistettujen tilastotietojen tarkentuminen eli revisio on osa normaalia tilastotuotantoa ja merkitsee laadun parantumista. Periaatteena on, että tilastotiedot perustuvat parhaaseen saatavilla olevaan aineistoon ja tietoon tilastoitavasta ilmiöstä. Toisaalta tarkentumisesta pyritään tiedottamaan mahdollisimman läpinäkyvästi ennakkoon. Ennakkoviestinnällä varmistetaan, että käyttäjät pystyvät varautumaan tietojen tarkentumiseen.

Tilastojulkistusten tietojen tarkentumisen taustalla on useimmiten aineiston täydentyminen. Tällöin uusi, revisioitutilastoluku perustuu laajempaa tietopohjaan ja kuvaa ilmiötä entistä tarkemmin.

Tilaston tietojen tarkentuminen voi liittyä myös käytettävään laskentamenetelmään, kuten lukujen vuosittaiseen täsmäytykseen tai painorakenteen päivitykseen. Myös perusvuoden ja käytettyjen luokitusten muutoksesta aiheutuu tietojen tarkentumista.

Suhdannetilastojen kausitasoitettut tiedot tarkentuvat käytettävän laskentamenetelmän vuoksi. Uuden aikasarjahavainnon tuoma lisäinformaatio hyödynnetään mallipohjaisessa laskentamenetelmässä, mikä heijastuu vanhojen julkistustenmuutoksena. Tuoreimpien kausitasoitettavien lukujen revisoitumisesta kerrotaan tilastojen julkistuksissa ja laatuselosteissa.

Keskeisten suhdannetilastojen ja joidenkin vuositilastojen yhteydessä julkistetaan myös taulukko toteutuneista revisioista. Taulukosta ilmenee, miten viimeisten tilastoajankohtien tiedot ovat muuttuneet ensimmäisen ja tuoreimmantilastojulkistuksen välillä.

Tilastoprosessi

Lähdeaineistot

Työvoimatutkimuksen tilasto perustuu kuukausittain haastattelemalla kerättävään otosaineistoon. Kuukausittainen otos on noin 12 500 henkeä ja tiedot kerätään puhelinhaastatteluilla ja osin käyntihaastatteluilla sekä vuodesta 2021 alkaen verkkokyselomakkeella. Samaa vastaajaa haastatellaan kaikkiaan viisi kertaa.

Kotitalousosan aineisto perustuu viidennen kerran vastaajien sekä heidän kotitalouksiensa jäsenten antamiin tietoihin. Kotitalouden jäsenten tiedot kerätään puhelinhaastatteluilla ja osin käyntihaastatteluilla sekä vuodesta 2022 alkaen verkkokyselomakkeella. Kotitalousosan vuosiaineisto sisältää noin 15 000 kotitalouden tietoja.

Aineiston laadinnassa on käytetty seuraavia hallinnollisia aineistoja ja tilastorekistereitä: Digi- ja väestöviraston väestötietojärjestelmä ja Tilastokeskuksen henkilö- ja asuntotietovaranto, Verohallinnon verotietokanta, Tilastokeskuksen tutkintorekisteri, Tilastokeskuksen yritysrekisteri.

Tiedonkeruun tiheys

Työvoimatutkimuksen tilasto perustuu jatkuvaan tiedonkeruuseen, jossa kaikki vuoden viikot ovat tutkimusviikkoja.

Tiedonkeruumenetelmä

Työvoimatutkimuksen kotitalousosan tiedot kerätään Tilastokeskuksen haastattelijoiden tekemillä tietokoneavusteisilla puhelinhaastatteluilla ja osin käyntihaastatteluilla sekä vuodesta 2022 alkaen verkkokyselomakkeella. Lomakkeen kielenä käytetään haastateltavan valinnan mukaan joko suomea, ruotsia tai englantia (tilastovuodesta 2014 lähtien).

Haastattelu tehdään pääsääntöisesti haastateltavalle itselleen. Jos haastateltavaa itseään ei ole mahdollista haastatella, pyritään haastattelutiedot kysymään ns. sijaisvastaajalta. Sijaisvastaajana voidaan käyttää haastateltavan puolisoa /avopuolisoa, vanhempia ja 15 vuotta täyttänyttä lasta tai sisarta / veljeä. Näiden tulee asua samassa kotitaloudessa kuin haastateltava.

Tutkimus on paneelitutkimus, jossa samaa henkilöä haastatellaan viisi kertaa. Viimeisellä kerralla kerätään myös kohdehenkilön kotitalouden jäsenten tiedot.

Suurimmassa osassa Euroopan maita työvoimatutkimuksen tiedonkeruu perustuu kotitalousotokseen, mikä tarkoittaa sitä, että haastatellaan samalla kertaa kaikki tietyssä

osoitteessa asuvat saman kotitalouden jäsenet. Suomen lisäksi mm. Ruotsissa, Tanskassa ja Sveitsissä otos on henkilöpohjainen, eli haastateltavana on vain otokseen poimittu kohdehenkilö. EU:n asetus kuitenkin edellyttää tietoja myös kotitaloudesta. Sen mukaan jäsenvaltiot, joiden otos on henkilöpohjainen, voivat toimittaa kyseisten kotitalouksien muita jäseniä koskevat tiedot riippumattomien havaintojen osaotoksesta ("kotitalouden osaotos"), joka on muodostettu siten, että kotitalouden osaotos käsittää vähintään yhden tutkimuskerran vuosineljännestä kohti tai vähintään 15 000 kotitaloutta. Suomessa tämä on siis ratkaistu niin, että viidennellä eli viimeisellä haastattelukerralla kerätään myös kohdehenkilön kotitalouden jäsenten tiedot.

Vuonna 2023 työvoimatutkimuksessa kotitalousosan kyselyyn oli saatu yli 33 000 henkilön vastauksia, oli haastateltu noin 15 000 kotitaloutta. Kaikista viidennellä tutkimuskerralla kuukausityövoimatutkimuksessa vastanneista noin 92 % vastasi myös kotitalousosaan.

Aineiston/datan validointi

Vastausaineisto ja myös lopullinen aineisto tarkistetaan aineiston muodostamisen ja raportoinnin yhteydessä. Suurin osa vastausaineistontarkistuksista on ohjelmoitu Blaise-sovelluksen kysymyslomakkeelle. Ohjelmassa on vastaustiedoille sallittujen arvojentarkistuksia sekä loogisuustarkistuksia. Käsittelyvaiheessa koodataan puuttuvat tai virheelliset toimiala- ja ammattikoodit käytetään apuna muita vastaustietoja. Myös tuntemattomiksi jääneet työtuntitiedot imputoidaan. Julkaistavan tilastotiedon laatu varmennetaan mm. katotarkasteluilla, vertailuilla työ- ja elinkeinoministeriön tietoihin ja vastaavan ajankohdan tuloksiin edellisenä vuonna. Tämän lisäksi Eurostat validoi ja tarkastaa neljänneksittäin saapuneenaineiston yksityiskohtaisesti.

Tiedon käsittely

Kotitalousosaotoksen vastaustiedot korotetaan sekä henkilö- että kotitaloustasolla edustamaan kohdeperusjoukkoa painokertoimien avulla. Painotus korjaa vastauskadon vaikutuksia, parantaa tietojen tilastollista tarkkuutta ja toteutetaan EU-asetuksen painotus- ja konsistenssivaatimusten mukaisesti. Siinä sovelletaan painojen kalibroitimenetelmää, jossa käytettyjen taustamuuttujien (ns. kalibroitimuuttujien) reunajakaumat kalibroidaan vastaamaan kohdeperusjoukon reunajakaumia eli väestörakennetta.

Henkilötasolla kotitalousosaotoksen painokerroin summautuu Suomen väkilukuun (ilman asevelvollisia) ja käytettäviä kalibroitimuuttujia ovat sukupuoli, ikäryhmä, koulutusaste ja äidinkieli. Lisäksi painotuksessa huomioidaan EU-asetuksen konsistenssivaatimukset, joiden mukaan osaotoksesta laskettavien henkilötason tulosten on oltava yhdenmukaisia henkilöotoksesta laskettujen vuositulosten kanssa seuraavissa

osajoukoissa: työlliset, työttömät ja työvoimaan kuulumattomat sukupuolittain ikäryhmissä 15-24, 25-34, 35-44, 45-54 ja 55+.

Kotitaloustasolla painokerroin kalibroidaan nk. benchmarking-menetelmällä vastaamaan Tilastokeskuksen tulonjakotilaston kotitalouksien rakennetta alueen (maakunnat ja pääkaupunkiseutu) ja kotitalouden henkilömäärän mukaan.

Vuoden 2021 alusta kotitalousaineiston painojen kalibroinnissa aputietona käytetään tulonjakotilastosta saatuja kotitalouskohtaisia tietoja:

- kotitalouksien lukumäärä
- kotitalouden koko sekä
- aluetieto (maakuntajako, jossa pääkaupunkiseutu erikseen)

Rekisteripohjaisia henkilötietoja:

- ikäryhmät
- sukupuoliryhmät
- kieli

Aputietoina kalibroinnissa käytetään työ- ja elinkeinoministeriön työnhakijarekisteriä, tilastokeskuksen tutkintorekisteriä ja verohallinnon tulorekisteriä. Myös tulosten yhdenmukaisuuden vuoksi käytetään kuukausittaisen työvoimatutkimuksen tietoja työllisistä, työttömistä ja työvoiman ulkopuolella olevista sukupuolen ja iän mukaan.

Kotitalousosaotoksen painojen kalibrointiin käytetään ranskalaista CALMAR2-ohjelmaa ja keskivirhe-estimointiin ruotsalaista ETOS-ohjelmaa.