

# Pohdintaa Suomen väestöennusteen oletuksista poikkeusvuonna 2024

Juha Alho\*

08/11/2024

---

\*J.A. on emeritusprofessori (Helsingin yliopisto) ja Senior Fellow (ETLA). Yhteystiedot: posti: juha.mikko.alho@gmail.com; tekstiviestit: +358400271332. Muistion tekemisestä sovittiin tutkijapalaverissa 29.10.2024. Kirjoittaja kiittää M. Rapoa Tilastokeskuksen ennusteen oletusten täsmentämisestä ja lähtötiedoista, N. Keilmania (Oslo yliopisto) pohjoismaisia ennusteita koskevista tiedoista sekä J. Lassilaa ja T. Valkosta (ETLA) hyödyllisistä keskusteluista. Muistio on osa Strategisen tutkimuksen neuvoston LIFECON-projektia, päätös 345218. Esitetyistä mielipiteistä vastaa yksin kirjoittaja.

## Tiivistelmä

Tilastokeskuksen väestöennusteen päivitys tapahtui poikkeuksellisen vaikeassa tilanteessa 2024. Mikä on COVID19:n vaikutus kuolevuuteen ja maahanmuuttoon ja jatkuuko vuonna 2010 alkanut syntyvyyden romahdus? Myös Ukrainan sodan ja muuttoliikkeen rakenteen muutoksen vaikutus maahanmuuttoon piti arvioida trendiennustetta varten. Kriittisessä tarkastelussa sekä kuolevuutta, syntyvyyttä että lähivuosien nettomaahanmuuttoa koskevat Tilastokeskuksen ratkaisut vaikuttavat perustelluilta. Tässä muistiossa on arvioitu tilastollisen aikasarja-analyysin keinoin nettomuuton trendin taso vuosille 2022–2023. Tätä voisi pitää trendiennusteessa myös pitkän ajan odotettuna tasona. Arvioksi saadaan 24 000, mikä on selvästi matalampi kuin päivityksessä käytetty 40 000. Tämä matalampikin taso on korkeampi kuin Norjan ja Ruotsin tuoreissa väestöennusteissa. Johtopäätöksenä on, että on syytä tuottaa vaihtoehtoinen päivitys, joka on muuten sama kuin Tilastokeskuksen 2024 päivitys, mutta pitkän ajan nettomaahanmuuton taso on 24 000.

The updating of the Statistics Finland population forecast occurred in an exceptionally difficult circumstance, in 2024. What would be the effect of COVID19 on mortality and immigration? Will the collapse of fertility that started in 2010, continue? Also the war in Ukraine, and the structural change in immigration, had to be assessed for the drawing up of a trend forecast. In a critical evaluation, the choices made by the forecasters regarding mortality, fertility, and the immigration in the first coming years appear well founded. Using tools of statistical time-series analysis, the level of net migration for 2022–2023 was estimated. This can be taken as the estimate of the level over the long term, in a trend forecast. The estimate is 24 000. This is lower than 40 000 that was used in the update. In a comparison to the recent forecasts in Norway and Sweden, we find that even this level is higher than that assumed in those countries. The conclusion is to produce an alternate update that replicates the assumptions of Statistics Finland as regards mortality, fertility, and short-term net migration, but assumes that the long-term level of net migration equals 24 000.

Avainsanat: epävarmuus, hedelmällisyys, kuolevuus, maahanmuutto, nettomuutto, trendiennuste.

Key words: fertility, immigration, mortality, net migration, trend forecast uncertainty.

# 1 Trendiennusteiden laadinnan periaatteista

Väestön määrä vuonna  $t + 1$  saadaan vähentämällä vuoden  $t$  väestöstä vuoden aikana kuolleet, lisäämällä syntyneet ja lisäämällä (positiivinen tai negatiivinen) nettomuutto. Laskut suoritetaan ikä- ja sukupuoliryhmittäin minkä takia menetelmää kutsutaan *kohorttikomponenttimenetelmäksi*<sup>1</sup>.

*Trendiennusteesta* puhutaan, kun halutaan korostaa sitä, että väestön muutoskomponentteja koskevat oletukset eivät perustu taloudellista käyttäytymistä tai sosiaalisten olosuhteiden muutoksia koskeviin arvioihin. Kyseessä on “*business-as-usual*”-oletus, ts. aiemman kehityksen oletetaan jatkuvan.<sup>2</sup> Tästä voi päättyä ajattelemaan, että ennusteen laskenta on mekaaninen suoritus.

Näin useasti onkin. Jos kuolevuudessa, hedelmällisyydessä ja muuttoliikkeessä ei tapahdu merkittäviä muutoksia, aiemman ennusteen päivitys on rutiinisuoritus. Yllätyksiä kuitenkin tapahtuu. Ne voivat olla lyhytaikaisia shokkeja tai pitkäaikaisen trendin muutoksia. Koska väestön muuttuminen tapahtuu “korkoa korolle” -periaatteella, vaikutus voi kumuloitua ajan kuluessa.

Ennusteen tekeminen vaatii aina asiantuntemusta, harkintaa ja valmiutta reagoida muuttuviin olosuhteisiin. Oletukset eivät saisi johtaa epäloogisuuksiin esimerkiksi ikäryhmien välillä, sukupuolten välillä tai maiden välillä. On tavallista tehdä vertailuja muihin maihin sen varmistamiseksi, ettei ennakoitu kehitys näytä oudolta kansainvälisessä tarkastelussa. Näennäisestä suora-  
viivaisuudestaan huolimatta trendiennuste pyritään tekemään niin, että siinä tiivistyy alan paras tietämys.<sup>3</sup>

Menneinä vuosikymmeninä oli tapana taittaa peistä siitä, onko trendiennuste “oikea ennuste” vai onko se vain “projektio”, jonka uskottavuuteen ennusteen laatija ei ota kantaa. Edellä sanotun perusteella keskustelu oli ja on turhaa. Ei ole mieltä julkaista laskelmia, joiden tiedetään jo ennustehetkellä antavan epäuskottavan kuvan tulevasta väestönkehityksestä. Virallisten ennusteiden käyttäjät ovat aina olleet oikeutettuja oletamaan, että tätä perusperiaatetta noudatettaisiin.<sup>4</sup> Tässä kirjoituksessa tarkastelen tuoreen

<sup>1</sup>Termi *kohortti* viittaa tässä kirjoituksessa tietynä vuonna syntyneisiin.

<sup>2</sup>Tämä lähestymistapa on osoittautunut hyödylliseksi siksi, että pysyviä yhteyksiä väestöilmidiöiden ja talouden välillä on ollut vaikea löytää. Silloinkin, kun yhteyden on arveltu olevan olemassa, ongelma on ollut syytekijöiden ennustaminen, ks. esim. Land (1986).

<sup>3</sup>Veneilyallegoria saattaa olla hyödyllinen. Avoimella selällä ja hyvissä olosuhteissa matkavene kulkee turvallisesti autopilotilla. Mutta kun tulee kapeikkoja vastaan ja tuuli nousee, on paras pitää kättä ruorilla. Jos uhkaa ajautuminen reamarin väärälle puolelle, voi äkkijarrutus olla paikallaan. Otsikon “poikkeusvuosi” viittaa viime mainittuun.

<sup>4</sup>“So long as numbers which purport to be possible future populations are published they will be regarded as forecasts or predictions, irrespective of what they are called by

Tilastokeskuksen ennusteen oletuksia tästä näkökulmasta.

Väestöä koskevien ennusteiden tarkkuutta on systemaattisesti tutkittu runsaan 30 vuoden ajan. On käynyt ilmi, että ennusteet ovat selvästi epävarmempia, kuin suuri yleisö tai edes väestötieteen asiantuntijat ovat uskooneet (ks. esim. Keilman, 1990). Yksi syy siihen, että tulevaa väestöä koskevien laskelmien ennusteluonnetta ei haluta korostaa, lienee tässä.

Järkevämpää olisi kuitenkin myöntää, että tietomme väestömuutosten syistä on toistuvasti osoittautunut erheelliseksi. Kun parhainakin pidetyt ennusteet ovat osoittautuneet enemmän tai vähemmän virheellisiksi, olisi järkevintä hyväksyä, että näin on luultavasti jatkossakin (vrt. Keyfitz, 1982). Kun epävarmuutta ei näytä olevan mahdollista eliminoida, järkevää olisi arvioida sen suuruutta ja käyttää tätä tietoa päätöksenteossa. Hyvän esimerkin tästä antaa osakemarkkinoiden volatiliiteetin<sup>5</sup> käyttö johdannaiskaupassa.<sup>6</sup>

Luvussa 2 käsitellään Tilastokeskuksen tuoreen ennusteen oletuksia. Nähdään, että erityisesti muuttoliikettä koskevat oletukset ovat pitkällä tähtäimellä ongelmallisia. Luvussa 3 tarkastellaan muuttoliikettä aikasarja-analyysin keinoin. Tuloksia verrataan tuoreisiin pohjoismaisiin ennusteisiin luvussa 4 ja luvussa 5 peilataan ennusteiden eroja suhteessa niiden odotettavaan epävarmuuteen. Luvussa 6 pohditaan parannusmahdollisuuksia.

Osa muistion luvuista sisältää tekniluontoista kieltä, joissa tehtyjä mallitusratkaisuja perustellaan. Nämä luvut on merkitty otsikon perässä olevalla tähdellä (\*). Päätuloksista kiinnostunut lukija voi harpata näiden yli.

## 2 Tilastokeskuksen ennusteen oletukset poikkeusvuonna 2024

Tilastokeskus julkaisi päivityksen väestöennusteeseensa 24.10.2024. Muutos edelliseen, 30.9.2021 julkaistuun ennusteeseen oli suuri. Mikä aiheutti sen?

Taustana oli, että vuonna 2020 alkanut koronaepidemia vaikutti kuolevuuteen ja kenties maahanmuuttoonkin. Hedelmällisyys oli jyrkästi laskenut vuoden 2010 jälkeen ja oli johtanut kaikkien aikojen matalimpaan tasoon. Ukrainan sota oli ajanut pakolaiset liikkeelle, Suomeenkin. Lisäksi Suomen maahan- ja maastamuutossa Aasian merkitys korostui.

Erityisesti nettomaahanmuuttoa koskevat oletukset olivat päivityksessä muuttuneet merkittävästi. Miten shokit näkyvät Tilastokeskuksen trendiennusteessa?

---

the demographers who prepare them.” (Dorn, 1950).

<sup>5</sup>Volatiliiteetti-termin suomennokseksi näyttäisi vakiintuvan *heilunta*.

<sup>6</sup>Alan tutkijoiden piirissä tämä johtopäätös on pitkään ollut yleisesti jaettu, ks. esim. Bijak et al. (2015).

## 2.1 Kuolevuus

Tilastokeskus on 1990-luvulta alkaen olettanut ennusteissaan, että kuolevuuden aleneminen ei ennusteperiodilla loppuisi. Tulokset ovat olleet hyviä. Ennusteissa on oletettu lähivuosina havaitun kuolevuuden laskun jatkuvan kussakin ikä- ja sukupuoliryhmässä aiempaa vauhtia.<sup>7</sup>

Kaksi esimerkkiä Tilastokeskuksen trendiennusteen 2024 hienovaraisuudesta ovat seuraavat. Ensinnäkin, kun alle 50-vuotiaiden joukossa kuolevuuden aleneminen on laskettu vertailemalla alenemista 5-vuotisperiodilta 1989–1993 periodille 2018–2023, niin näitä vanhempien kohdalla aloitetaan vasta periodista 1999–2003. Tämä on perusteltua, sillä kuolevuuden alenemisen on havaittu hitaasti siirtyvän nuoremmista ikäryhmistä vanhempiin.

Toiseksi, vaikka COVID19 ei juuri vaikuttanut kuolevuuteen varsinaisina koronavuosina 2020–2021, se yllätti vuonna 2022. Viisaasti, vuosi 2022 on jätetty pois kuolevuuden alenevan trendin arvioinnista, sillä jo vuonna 2023 kuolevuus jälleen laski.

## 2.2 Hedelmällisyys

Vuosien varrella Tilastokeskus lakkasi “hienosäätämästä” hedelmällisyyden trendiennusteita. Ennuste kokonaishedelmällisyydelle on yksinkertaisesti tuorein havaittu arvo.<sup>8</sup>

Voidaan kuitenkin perustellusti kysyä, miksi ei käytetty 5-vuotiskeskisarvoa niin kuin kuolevuudessa. Samoin kuin kuolleiden lukumääriin, liittyy synty-miin vuotuista vaihtelua, joka ei ole osa trendinmuutoksia. Kun ennusteessa käytetty kokonaishedelmällisyyden arvo 1.26 oli kaikkien aikojen matalin ja verrokkimaiden matalin, niin olisiko syy havaittuun matalaan arvoon voinut osin olla vuotuisessa vaihtelussa?

Niinpä vaihtoehtoinen arvio olisi voinut olla 1.35, vuosien 2019-2023 keskiarvo. Tätä arvoa käyttäen lasten lukumäärät olisivat runsaat 7 % suurempia joka vuosi. Kun vuonna 2023 syntyi 43 383 lasta, ero vastaa 3 100 lasta vuodessa. Tätä suuruusluokkaa oleva lisäys tulisi siis kaikkiin tulevi-na vuosina syntyviin kohortteihin ja se kasvattaisi väestön kokonaismäärän ennustetta 30 vuoden päästä noin 100 000:lla.

Ero ei kuitenkaan ole olennainen, jos sitä vertaa esimerkiksi vuosimuutosten kokoon. Tilastokeskuksen syntyvyysennustetta voikin pitää hyvin perusteltuna.

---

<sup>7</sup>Vanhusväestön määrä on edelleen näyttänyt tulevan hieman aliarvioitua. Mm. tästä syystä tapaturmavakuutuslaitoksissa tarvitaan tarkempia analyysejä, ks. esim. TVK (2021).

<sup>8</sup>Tämä on tilastollisesti optimaalinen ennuste, jos aikasarja on ns. martingali.

## 2.3 Maahanmuutto

Hankalimmaksi muutoskomponenteista osoittautui maahanmuutto. Ukrainan sota johti suureen pakolaisvirtaan kaikkialle Eurooppaan, ml. Suomeen.

On syytä huomata, että muuttoliikkeen muutoksilla ja niiden syillä ei aina ole yksiselitteistä suhdetta. Esimerkiksi kun Suomen “suuri muutto” kaupunkeihin ja Ruotsiin tapahtui vuoden 1970 tienoilla, niin syyksi voidaan osoittaa talouden rakennemuutos: muutettiin maaseudulta kaupunkeihin kiihokkeena korkeampi elintaso teollisuudessa ja palvelualoilla. Mutta nämä tekijät olivat vaikuttaneet taustalla pitkään. Kyse on myös dynaamisesta prosessista, johon vaikuttaa *innovaatioiden diffuusio* (Rogers, 2003): tiedot muuttajien hyvistä (tai huonoista) kokemuksista olivat yksilöille tärkeitä tekijöitä muuttopäätöstä tehtäessä. Muuton ajoitus on siten osin riippumaton sen juurisyiden kehityksestä.

Monet sitten palasivat takaisin Ruotsista todennäköisesti samanlaisten tekijöiden ajamana tai houkuttelemana – joskaan eivät enää maaseudulle – ja muuttoaalto päättyi. Muuttoliike voi siis suhdannevaihteluiden aiheuttamien aaltojen lisäksi olla luonteeltaan “purskeista”.

Tilastokeskuksessa tunnistettiin shokkien olemassaolo vuosina 2021–2023 ja ne pyrittiin ottamaan ennusteessa huomioon siten, että shokkien aiheuttaman tulovirran oletettiin joitain vuosia jatkuvan ja sitten laantuvan. Ratkaisu on luonteva. Purskeen kestoa koskevasta oletuksesta voi toki keskustella, mutta kukaanhan ei varmuudella tiedä, miten esimerkiksi Ukrainan sota tulee jatkumaan.

Ongelmaksi jäi kuitenkin se, että shokin jälkeen nettomuuton oletettiin jäävän pysyvästi tasolle 40 000 vuodessa. Perusteena oletukselle esitettiin, että tämä taso riittäisi suurin piirtein pitämään väestön koon ennallaan. Lisäksi mainittiin, että muuttovoitto itäisestä Aasiasta oli viime vuosina merkittävästi kasvanut.

ETLA:ssa oli jo aiemmin arvioitu (Alho, Kangasharju, Lassila ja Valkonen, 2023), että työikäisen väestön koon pitäminen vakiotasolla edellyttäisi 44 000:n henkilön nettomaahanmuuttoa vuosittain. Kyseessä ei ollut ennuste vaan laskuharjoitus siitä, mitä tarvittaisiin, jos vakiotyövoimaan pyrittäisiin. Itä-Aasian maista saatu muuttovoitto saattoi hyvin olla tietoisien rekrytointipolitiikan ansiota yhdessä tai erikseen lähtömaiden olosuhteiden muutosten kanssa (vrt. Lähdemäki ja Rahko, 2024). Tilastollisessa mielessä kyseessä oli kuitenkin jyrkkä poikkeama Aasiasta tulevan maahanmuuton volyymissa, eli *oudokki*<sup>9</sup>.

Kun edellisessä Tilastokeskuksen ennusteessa nettomaahanmuutto oli 15 000 vuodessa eli ennustevuotta edeltävä arvo, niin voidaan kysyä, miten voi-

---

<sup>9</sup>Tämä hauska sana on nopeasti vakiintunut termin *outlier* suomennokseksi.

daan perustella määrän kasvattamista 2.7-kertaiseksi. Lisäksi koska muuttajat ovat tyypillisesti hedelmällisyydessä, positiivinen nettomuutto kasvattaisi myös syntyvyyttä korkoa korolle -periaatteella.

## 2.4 Vertailun vuoksi vuosi 1974

Otsikon mukaan ennusteen päivitys vuonna 2024 tapahtui poikkeuksellisissa olosuhteissa. Tuorein verrannollinen oudokkivuosi oli 1974. Mitä silloin tapahtui?

Vuosina 1963–1973 kokonaishedelmällisyys laski arvosta 2.66 lasta naista kohti arvoon 1.50, eli 44 %. Vuosina 1963–1970 tuli muuttotappiota (lähinnä Ruotsiin muuton takia) yhteensä 158 544, tästä 76 880 vuosina 1969–1970. Kun vuosina 1971–1973 paluumuutto vahvistui, saatiin näinä vuosina muuttovoittoa 12 325. Kuolevuuden lasku jatkui noina vuosina, eikä siinä nähty ongelmaa.

Vuoden 1974 ennusteessa (Tilastokeskus, 1974) oletettiin, ettei hedelmällisyys enää laskisi, vaan pysyisi suurin piirtein ennallaan. Hedelmällisyyden lasku pysähtyikin ja nousi noin 0.3:lla lapsella naista kohti. Ennuste osui tältä osin oikeaan. Ajalle tyypillisesti oletettiin kuolevuuden pysyvän ennallaan tai jonkin aikaa laskevan ja sitten jäävän saavutetulle tasolle. Muuttoliikeoletus oli tulevaan kehitykseen verraten, erityisesti myöhempinä ennustevuosina, pessimistinen.

Oletusten seurauksena väestön koko tuli aliarvioiduksi, niin että vuoden 2000 keskiennuste oli 4.65 miljoonaa kun tuolloin väestön koko oli 5.18 miljoonaa.

## 3 Maahanmuuttoshokki aikasarja-analyysin valossa

Shokkien suuruutta voidaan arvioida laatimalla edeltävän aineiston perusteella ennuste shokkivuosille. Toteutuneen arvon ja ennusteen ero voidaan tulkita shokin aiheuttamaksi. Tähän arvioon, niinkuin kaikkiin muihinkin ennusteisiin liittyy oma epävarmuutensa. Sen suuruus voidaan määrittää, kun käytetään aineistoon sopivaa aikasarjamallia.

Tilastokeskuksen StatFin-tietokannasta saadaan nettomaahanmuuttotiedot vuosilta 1990–2023 sukupuolen, lähtömaan ja kansalaisuuden mukaan. Shokkivuosina pidetään vuosia 2022–2023 ja mallit laaditaan vuosien 1990–2021 perusteella.

Lähetettäjämaat luokitellaan kolmeen osaan: Ukraina, Aasia ja Muut maat. Sukupuoli on potentiaalisesti tärkeä, sillä muutossa Ukrainasta sotaa paen-

neiden naisten rooli korostuu, kun yleensä nuoret miehet ovat perheen ensimmäisiä muuttajia. Kansalaisuus voisi puolestaan olla merkityksellinen COVID19:n roolin selvittämisessä. Se jaotellaan kahteen luokkaan: Suomi ja muut maat. Näin saadaan  $3 \times 2 \times 2 = 12$  aikasarjaa.

Etukäteen ajatellen lähtömaa on tärkein tekijä, joten analyysi suoritettiin näille kolmelle lähtöalueelle erikseen ja tutkittiin sukupuolen ja kansallisuuden roolia niiden sisällä.

### 3.1 Sovellettu aikasarja-analyysi\*

Tilastollisen ennusteteknologian kehittymisen myötä on mahdollista antaa tietokoneen etsiä aineistoihin parhaiten sopivat vähäparametriset ARIMA-mallit (Box, Jenkins, Reinsel ja Ljung, 2015) kullekin 12 aikasarjasta.<sup>10</sup> Koska havaintoväli 1990-2021 on varsin lyhyt, mallien identifiointia ei voida luotettavasti tehdä, ts. on monia malleja jotka sopivat lähes yhtä hyvin aineistoon.

Automatisoidun proseduurin tuottamia parhaiten sopivia malleja käytetään kuitenkin pelkästään trendiennusteiden tekoon kahdelle tulevalle periodille 2022-2023, jolloin niiden pitemmän aikavälin pätevyys on toissijainen kysymys.

### 3.2 Ukraina\*

Suomen ja Ukrainan välisessä muuttoliikkeessä Suomen kansalaisten osuus oli mitätön, joten muuttajien kansalaisuus voidaan kokonaan sivuuttaa analyysissä. Sukupuolella oli havaittava vaikutus. Vuonna 2020 oli pieni kasvu miesten nettomuutossa, jota ei ollut naisilla, mutta vuosina 2021-2022 naisten muutto ylitti odotusten mukaisesti selvästi miesten muuton.

Ukrainan nettomuuttoa voidaan analysoida tekemättä eroa kansallisuuden mukaan. Sukupuolen mukaiselle tarkastelulle on perusteita, mutta tämän kirjoituksen tarkoituksia varten sukupuolten käsittely yhdessä tuottaa yksinkertaisemman kuvan shokkivaikutuksesta. Molemmat kuitenkin tehtiin sen osoittamiseksi, millaisia eroja tämän tyyppisissä analyysissä voi syntyä.

Ukrainasta Suomeen kohdistuvan nettomuuton malliksi valikoitui miehille ARIMA(0,2,2), naisille ARIMA(0,2,1) ja yhteiselle sarjalle ARIMA(0,2,2). Niiden tuottamat ennusteet vuodelle 2022 ovat seuraavat (80 %:n ennustevälit suluissa<sup>11</sup>): miehet (M) 1 117 (1 016, 1 217), naiset (N) 569 (517, 622),

<sup>10</sup>Tässä käytettiin R. Hyndmanin **R**-ympäristöön kirjoittaman *forecast* -kirjaston *auto.arima* -funktiota. Mallin valinta perustuu ns. Akaiken informaatiokriteeriin (AIC) (Hyndman ja Khandakar, 2008).

<sup>11</sup>Tulkinta on että, ennusteen laatija olisi (periaatteessa) valmis lyömään 4 euroa vetoa yhtä vastaan sen puolesta, että ennusteväli sisältää tulevan arvon. Muttei enempää.

yhteinen sarja (Y) 1 722 (1 569, 1 875). Nähdään, että erikseen laskettujen ennusteiden summa on  $1\,117 + 569 = 1\,686$ . Ero  $1\,722 - 1\,686 = 36$  on merkityksetön. Vuoden 2023 ennusteet ovat vastaavasti: M 1 538 (1 377, 1 699); N 619 (538, 701); Y 2 250 (2 001, 2 498). Miesten ja naisten ennusteiden summa  $1\,538 + 619 = 2\,157$  on lähellä yhteisen sarjan ennustetta 2 250. Ero on pieni erityisesti kun sitä verrataan ennustevälien pituuteen. Tästä syystä ja yksinkertaisuuden vuoksi jatkossa käytetään yhteisen sarjan ennusteita. Vuosien 2022 ja 2023 toteutunut nettomuutto oli 1 488 ja 19 531. Nähdään, että vuoden 2022 ennuste 1 722 oli *korkeampi* kuin toteutuma 1 486. Tämä johtunee viiveestä, joka on kulunut väliaikaisen suojelun käynnistämisestä ja siihen liittyvästä kotipaikan myönnöstä.

Sodan vaikutus näkyi siis vasta vuonna 2023. Sen suuruus oli  $19\,531 - 2\,250 = 17\,281$ . Tällöin luku 2 250 tulkitaan sotaa edeltäväksi nettomaahanmuutoksi Ukrainasta, jonka voi trendiennusteessa olettaa jatkuvan sodan vaikutusten päätyttyä.

### 3.3 Aasia\*

Suomen ja Aasian välisestä nettomuutosta noin 3 % koski Suomen kansalaisia. Määrä on niin pieni, ettei sen tuomia eroja kannata tämän kirjoituksen tarkoitusten valossa ottaa huomioon. Sukupuolten välillä on selkeä ero vuonna 2016–2017. Tuolloin miehiä tuli yhteensä viitisen tuhatta enemmän kuin edellisinä tai myöhempinä vuosina. Muuten miesten ja naisten nettomuutto oli hyvin samantasoista, naisia oli hieman enemmän vuosina 1993–2007, miehiä vuosina 2016–2023. Tästä syystä Aasian kohdalla ei kannata myöskään tehdä eroa miesten ja naisten välillä. Näin ollen analyysit raportoidaan vain nettomuuton kokonaismäärille, jossa miehet ja naiset ovat yhdessä (= Y).

Parhaiten sopiva malli sarjalle Y on ARIMA(0,1,1) suuntatermillä<sup>12</sup>. Vuodelle 2022 ennuste on: Y 7 724 (6 255, 9 193). Vuodelle 2023 se on: Y 7 933 (6 291, 9 576). Havaitut arvot olivat 15 681 ja 22 650. Näiden erotus on  $15\,681 - 7\,724 = 7\,957$  vuonna 2022 ja  $22\,650 - 7\,933 = 14\,717$  vuonna 2023 arvioidaan siis shokin aiheuttamiksi.

Vaihtoehtoisesti kun ennustetta toistetaan eri ajankohtina, neljässä tapauksessa viidestä ennusteväli sisältää tulevan arvon. Sekä vedonlyönti- että frekvenssitulkinta ovat tarpeen.

<sup>12</sup>Tämä näyttää vakiintuvan käännökseksi termille *drift*.

### 3.4 Muut maat\*

Kun Ukraina ja Aasia otetaan erikseen, jäljelle jäävät Muut maat ovat perin heterogeeninen ryhmä.<sup>13</sup> Tällä “jäännösryhmällä” on kuitenkin se erityinen piirre, että Suomen kansalaisten muutto kohdistuu valtaosin niihin. Kun analysoidaan suomalaisten nettomuuttoa tämän maaryhmän suhteen, osoitetaan, ettei se olennaisesti muuta arviota muuttoshokin suuruudesta.<sup>14</sup>

Miesten ja naisten nettomaahanmuuton trendit olivat hyvin samanlaista. Naisten osuus oli noin 70 % miesten määrästä. Tästä syystä sukupuolta ei kannata ottaa huomioon. Parhaiten sopiva malli sarjalle Y on ARIMA(0,1,0). Vuodelle 2022 ennuste on: Y 13 665 (10 819, 16 510). Vuodelle 2023 se on: Y 13 665 (9 641, 17 689). Havaitut arvot olivat 17 194 ja 15 733. Näiden erotus on  $17\,194 - 13\,665 = 3\,529$  vuonna 2022. Vuonna 2023 se on  $15\,733 - 13\,665 = 2\,068$ . Nämä olisivat siis arviot shokin suuruudesta po. vuosina, ja 13 665 olisi sopiva arvo trendiennusteelle shokin jälkeen.

### 3.5 Yhteenveto

Edellä johdetut arviot maahanmuuton “normaalitasosta” vuodelle 2023 ovat yhteensä  $2\,250 + 7\,933 + 13\,665 = 23\,848$ .

Kun edellä kuvattu analyysi toistetaan ottamatta mitenkään huomioon sen enempää sukupuolta tai lähtömaata, eli sovitetaan malli nettomaahanmuuttoon sellaisenaan saadaan malliksi ARIMA(0,1,0) ja sen tuottama ennuste vuodelle on 22 393 (17 689, 27 097).

Muutamilla lisälaskelmilla saadaan parhaana pitämämme mallin ennusteen 23 848 ympärille 80 %:n ennusteväli (19 494, 28 201). Tämä on muistion keskeisin tulos.

Parhaan mallin ennuste on siis hieman optimistisempi nettomaahanmuuton suhteen kuin yksinkertaisin järkevä malli. Mallissa on enemmän parametreja kuin yksinkertaisemmassa mallissa, joten se sopii paremmin, ja ennusteväli on siksi kapeampi.

Edellä esitetyt analyysit eivät tietenkään ole tyhjentyviä. Kuten esimerkiksi Aasian kohdalla mainittiin, sieltä löytyy eri tavoin käyttäytyviä maaryhmiä. Toistuva havainto ennusteiden tarkkuuden arvioinnissa on kuitenkin ollut, että yksinkertaiset yhteiskunnallisiin ilmiöihin liittyvät ennusteet ovat

<sup>13</sup>Maita, joista Suomeen muutto on hyppäksenomaisesti kasvanut vuosina 2022–2023 ovat Bangladesh, Nepal, Pakistan ja Sri Lanka. Maita, joista muutto on kasvanut voimakkaasti, mutta tasaisemmin ovat Filippiinit, Intia, Iran, Kiina, Thaimaa ja Vietnam. Muuton rakenne on merkittävästi muuttunut.

<sup>14</sup>Menemättä yksityiskohtiin voidaan todeta, että Suomen kansalaisten *maastamuutto* on lähes jokaisena tarkasteluvuotena ollut suurempaa kuin maahanmuutto sekä miehillä että erityisesti naisilla.

usein lähes yhtä hyviä – joskus parempia – kuin monimutkaisiin kausaalimalleihin perustuvat ennusteet. Syynä on se, että kausaalimallien rakenne voi olla virheellinen ja/tai mallien parametrit saattavat muuttua ajan yli.<sup>15</sup>

## 4 Vertailu Norjaan ja Ruotsiin

Suomessa nettomaahanmuutto on ollut positiivista vasta vuodesta 1981. Noin vuoteen 2004 asti muuttovoitto oli 5 000:n luokkaa vuosittain. Väestöön suhteutettuna tämä oli EU:n matalimpia arvoja. Muut Pohjoismaat, erityisesti Ruotsi, ovat olleet muuttovoittomaita pitempään. Tätä taustaa vasten on informatiivista tarkastella, millaisia oletuksia niiden tuoreissa väestöennusteissa on tehty maahanmuutosta.

Norjassa ja Ruotsissa väestöennusteiden laadintaan nivotaan selvästi Suomea enemmän joko suoranaista ekonometrista mallitusta tai siihen verrattavaa skenaarioajattelua. Niissä siis tunnustetaan, että kyse on nimenomaan “ennusteista” eikä vain laskelmista.

Seuraavassa tyydytään kuvailemaan mitä näissä maissa nettomaahanmuutosta oletetaan. Molemmissa Ukrainan vaikutuksia kuvataan samaan tapaan kuin Suomessa: todetaan shokin olemassaolo ja oletetaan, että vaikutus häviää muutamassa vuodessa.

Norjan ennusteessa vuosille 2024–1970 (Thomas ja Tømmerås, 2024) nettomaahanmuuton ennuste laaditaan maahanmuuton ja maastamuuton ennusteiden erotuksena. Raportin kuvion 7.25 nettomuutto laskee jyrkästi vuoteen 2025 mennessä noin tasolle 17 000, jonka jälkeen se tasaisesti *laskee* noin 12 000:een vuoteen 2070 mennessä. Hienoista laskua perustellaan kahdella syyllä. Ensiksi otaksutaan että lähtömaiden taloudellinen tuottavuus kasvaa nopeammin kuin Norjan. Toiseksi, ennakoidaan, että lähtömaissa väestö alkaa ikääntyä, jonka seurauksena maastamuuttointensiteetti laskee. Ottamatta kantaa ensin mainittuun otaksumaan, jälkimmäinen on eittämättä totta ja vaikuttaa myös Suomeen.

Ruotsin ennusteessa vuosille 2024–2070 (SCB, 2024) nettomaahanmuutto lasketaan niinkään erotuksena. Sen ennakoidaan kasvavan vuoden 2024 arvosta 23 500 vuoteen 2070 mennessä tasolle 30 700. Kun Norjassa lähtöalueita on kolme<sup>16</sup> Ruotsissa tarkastellaan muuttoa syntymämaan mukaan. Niitä on viisi.<sup>17</sup> Nähdään, että kovin hienojakoista tarkastelua ei ole yritetty

<sup>15</sup>Esimerkiksi kuolevuuden ennusteita ei ole onnistuttu parantamaan ottamalla kuolinsyy huomioon, vaikka toisin voisi odottaa.

<sup>16</sup>I. Pohjoismaat, “vanhan EU:n” ja EFTA:n maat sekä U.S.A., Kanada, Australia ja Uusi-Seelanti; II “uudet” Itä-Euroopan EU-maat; III. Muu maailma.

<sup>17</sup>I. Ruotsi; II. Pohjoismaat ja “vanha EU”; III. Afrikka; IV. Aasia; V. Muut maat.

kummassakaan maassa.

## 5 Ennusteen epävarmuudesta\*

Luvussa 3.5 havaittiin, että käyttökelpoinen malli nettomaahanmuutolle on yksinkertainen satunnaiskulku (*random walk*). Sen paras ennuste on aina viimeisin arvo, tässä tapauksessa 22 400, ts. hieman alempi kuin parhaana pitämämme ennuste, joka oli 24 000. Satunnaiskulkumallin 80 %:n ennusteväli vuodelle 2023 on samalla esitystarkkuudella (17 700, 27 100). Toisin sanoen  $\pm 5000$  ennusteen ylä- ja alapuolella.<sup>18</sup>

Luvun 3.5 arvoja pyöristäen paras arvio pitkän ajan nettomuutolle on (hieman alle) 24 000, ja sen 80 %:n ennusteväli on (19 500, 28 200). Tämän tarkempaa arviota ei voida tekemiemme analyysien perusteella antaa.<sup>19</sup>

## 6 Päätelmät

Tilastokeskuksen väestöennuste toimii pohjana useiden hallinnonalojen (esim. VM ja ETK) julkista taloutta koskeville laskelmille. Vaikka ennusteiden yksityiskohdista voi olla erimielisyyksiä, yleisesti hyväksytään, että tarvitaan “yhdet luvut”, jotta laskelmien pohjaoletukset olisivat vertailukelpoisia.

Tilastokeskuksen ennuste perustuu huolelliseen harkintaan ennustekomponenttien (kuolleisuus, syntyvyys ja muuttoliike) viimeaikaisesta kehityksestä ja oletuksesta, että kehitys jatkuu. Ennusteen laskenta on yksinkertaista aritmetiikkaa ja sen oletukset ovat läpinäkyviä, mikä tukee tulosten uskottavuutta ja ymmärrettävyyttä.

Vuoden 2024 päivitystä tehdessään ennusteen tekijät ovat joutuneet poikkeuksellisen hankalaan tilanteeseen. Hyvin perustein laskelmissa on otettu huomioon, että Ukrainan sota on aiheuttanut muuttoshokin, joka ajan kuluessa väistyy. Sen sijaan pitkän aikavälin oletus 40 000:n muuttovoitosta tuntuu korkealta. Oletuksen perustelut ovat niinkään olleet niukat. Koska oletuksella on merkittävä vaikutus esimerkiksi julkista taloutta koskeviin kesävyyslaskelmiin, on ollut tarpeen selvittää tarkemmin, mistä on kyse.

Käyttämällä yksinkertaisia aikasarjamalleja voidaan laskea ennusteet vuosien 2022-2023 nettomaahanmuutolle, mikäli shokkeja ei olisi ollut. Riippumatta siitä, millä resoluutiolla analyysi tehdään, tulos näyttää olevan, että

<sup>18</sup>Tämä vastaa normaalijakaumaoletuksen perusteella keskihajontaa 3 900.

<sup>19</sup>Epävarmuusarvio voidaan koodata niin, että se voidaan viedä koko maan väestöä koskevaan stokastiseen ennusteeseen. Tätä varten on olemassa valmis ohjelma PEP, jonka piirteitä kuvataan esimerkiksi monografiassa Alho ja Spencer (2005).

perusteltu luku pitkän tähtäimen nettomuutolle olisi noin 24 000. Tämä on 60 % korkeampi kuin edellisen ennusteen arvo 15 000, mutta merkittävästi vähemmän kuin vuoden 2024 päivityksessä valittu arvo 40 000.

Vertailu Norjan ja Ruotsin tuoreisiin ennusteisiin osoittaa, että tämäkin arvo on korkeahko, kun otetaan huomioon maiden koko (vuoden 2023 lopussa Suomessa oli 5.6, Norjassa 5.6 ja Ruotsissa 10.6 miljoonaa asukasta). Toisin sanoen jos Suomen nettomuuttovoitoksi oletetaan noin 24 000/vuosi, niin se on osapuulleen kaksi kertaa niin korkea kuin Norjassa ja Ruotsissa oletettu taso. Kuitenkin sen käyttö trendiennusteessa näyttää selvästi perustellulta.<sup>20</sup>

Konkreettisenä ratkaisuna olisi, että tuotettaisiin uusi ennusteversio, joka olisi muuten sama kuin Tilastokeskuksen 2024 päivitys, mutta pitkän ajan maahanmuuton tasoksi valittaisiin 24 000.<sup>21</sup>

Maahanmuuttajien kotoutuminen on keskeinen yhteiskuntapoliittinen ongelma. Tämän muistion aiheen kannalta asiaan liittyy byrokraattiselta kuulostava, mutta tärkeä ongelma. Kun kotikuntastatuksen myöntämiseen liittyy se, miten kauan myöntöprosessit kestävät, niin vuosittaisiin tilastoihin voi tulla merkittäviä heilahduksia hallinnollisista syistä. Tiedossa ei ole miten tämä on vaikuttanut edellä käsiteltyihin nettomuuttotilastoihin.

Lopuksi kiinnitän huomiota erääseen aiheeseen liittyvään ongelmaan, joka voi muodostua tärkeäksi ajanoloon. Suomen väestötilastot ja po. ennuste koskevat *vakinaisesti maassa asuvaa väestöä*. Jos kotipaikan saanut maahanmuuttaja kuitenkin lähtee pysyvästi pois maastamme, tästä ei välttämättä saada tietoa, jollei muutto kohdistu johonkin Pohjoismaista. (Näistä tieto muutosta saadaan viranomaisyhteistyön perusteella.) Aiemmin tällaisia henkilöitä oli vain vähän, koska iso osa muutosta kohdistui Ruotsiin. Nyt kun muuttajia tulee muualta, ongelmasta voi tulla merkittävä. Silloin kun maasta poistumista ei rekisteröidä, syntyy virhe sekä muuttotilastoihin että maan väkilukuun. Tiedossani on, että tätä on havahduttu pohtimaan ainakin Norjassa.

## Viitteet

Alho J., Kangasharju A., Lassila J. ja Valkonen T. (toim.) (2023). Maahanmuutto ja työvoiman riittävyys – taloudellisten vaikutusten arviointia. *ETLA Raportti No 132*.

<sup>20</sup>Raportin Alho, Kangasharju, Lassila ja Valkonen (2023) kuvioista 6c saa karkean (mutta hieman optimistisen) käsityksen siitä, millaisia tuloksia olisi odotettavissa.

<sup>21</sup>Omia kestävyyslaskelmiamme varten tullaan tuottamaan edellä mainituilla periaatteilla toteutettu stokastinen ennuste marraskuun lopulla 2024.

- Alho J.M., Spencer B.D. (2005) *Statistical Demography and Forecasting*. New York: Springer.
- Bijak J., Alberts I., Alho J., Bryant J., Buettner T., Falkingham J., Forster J.J., Gerland P., King T., Onorante L., Keilman N., O'Hagan A., Owens D., Raftery A., Ševčíková H., Smith P.W.F. (2015) Probabilistic Population Forecasts for Informed Decision Making: Letter to the Editor. *Journal of Official Statistics* 31, 537-544.
- Box G.E.P., Jenkins G.M., Reinsel G.C. ja Ljung G.M. (2015) *Time Series Analysis: Forecasting and Control*, 5th Edition. New York: Wiley.
- Dorn H. E. (1950) Pitfalls to population forecasts and projections. *Journal of the American Statistical Association* 45, 311-334.
- Hyndman R.J. ja Khandakar Y. (2008) Automatic time-series forecasting: the forecast package for **R**. *Journal of Statistical Software* 27, issue 3.
- Keilman N. (1990) Uncertainty in national population forecasting: Issues, backgrounds, analyses, recommendations. Swets & Zeitlinger.
- Keyfitz N. (1982) Can knowledge improve forecasts? *Population and Development Review* 8, 729-751.
- Land K. (1986) Methods for national population forecasting: A review. *Journal of the American Statistical Association* 81, 888-901.
- Lähdemäki S. ja Rahko J. (2024) Aasialaisten maahanmuutto Suomeen on kasvanut merkittävästi. *ETLA Erikoisartikkeli* 9 (4.11.2024).
- Rogers E. (2003). *Diffusion of Innovations* (5th ed.). New York: Simon and Schuster.
- Statistiska Centralbyrån. Sveriges framtida befolkning 2024-2070. *Demografiska rapporter* 2024:1.
- Tapaturmavakuutuskeskus (2021) Työtaturma-, liikenne- ja potilasvakuutuksen referenssikuolevuusmalli K2021. *Raportti* 21.8.2021.
- Thomas M.J. ja Tømmerås A.N. (2024) Norway's 2024 national population projections. Results, methods, and assumptions. Statistics Norway. *Reports* 2024/19.
- Tilastokeskus (1974) Väestöennusteet 1973–2000. *Tilastollisia tiedonantoja* N:o 52. Helsinki: Tilastokeskus.