

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Juvonen, Petteri; Anttonen, Jetro; Fornaro, Paolo; Nissilä, Wilma; Nyberg, Henri; Pönkä, Harri

Title: Aikasarjamallit apuna Suomen talouden seurannassa

Year: 2019

Version: Published version

Copyright: © 2019 Kansantaloudellinen yhdistys

Rights: In Copyright

Rights url: <http://rightsstatements.org/page/InC/1.0/?language=en>

Please cite the original version:

Juvonen, P., Anttonen, J., Fornaro, P., Nissilä, W., Nyberg, H., & Pönkä, H. (2019). Aikasarjamallit apuna Suomen talouden seurannassa. *Kansantaloudellinen aikakauskirja*, 115(3), 440-457.

https://www.taloustieteellinenyhdistys.fi/wp-content/uploads/2019/10/KAK_3_2019_nettiin-44-61.pdf

Aikasarjamallit apuna Suomen talouden seurannassa

**Petteri Juvonen, Jetro Anttonen, Paolo Fornaro,
Wilma Nissilä, Henri Nyberg ja Harri Pönkä**

Viimeisten vuosikymmenien aikana kansainvälisessä ekonometrisessä tutkimuskirjallisuudessa on esitetty useita makrotaloudellista tilaa kuvaavien muuttujien informaatiota yhdistäviä lyhyen aikavälin mallinnus- ja enustemenetelmiä. Näitä ns. nowcasting-menetelmiä on myös onnistuneesti hyödynnetty ja sovellettu Suomen talouden seurantaan. Tässä artikkelissa esittelemme katsauksen monella taholla tehtyyn kehitystyöhön ja näiden bankkeiden yhteydessä saatuihin tuloksiin Suomen aineiston tapauksessa. Suomen taloutta koskevien subdanneindeksien hyödyntämisen myötä subdanteiden käänne pisteiden määrittäminen on tarkempaa ja käänne pisteiden tuottamia taantumajaksoja voidaan vastaavasti ennustaa binääriavastemalleja käyttäen. Suomen Pankin nowcasting-malli mahdollistaa puolestaan uusien tilastojulkistusten uutisarvon analyysin. Tilastokeskuksessa ja Etlassa on vastaavasti hyödynnetty moderneja koneoppimisen menetelmiä, jotta puutteellisesta mikroaineistosta kyetään tuottamaan bruttokansantuotteen pikaestimaatteja aiempaa lyhyemmällä viiveellä. ETLAnow-projektissa hyödynnetään puolestaan mm. uusia Google-hakutilastoja työttömyyden ennustamisessa.

Ajantasaisen ja samalla luotettavan tiedon saaminen talouden tilasta on keskeinen haaste talouden päätöksentekijöille. Neljännesvuositainen reaalin bruttokansantuote on seuratuin kokonaistaloudellinen muuttuja. Sen on-

gelmana on kuitenkin pitkä julkaisuviive ja siihen myöhemmin tehty, usein suuretkin jälkikorjaukset. Talouden tilaa – kokonaistaloudellista aktiviteettia – kuvaavien muuttujien tilastotietojen reaaliaikaisuuteen liittyvien seik-

KTT Petteri Juvonen (petteri.juvonen@bof.fi) toimii tutkijatohtorina Helsingin yliopistossa (Helsinki Graduate School of Economics) ja Jyväskylän yliopiston kauppakorkeakoulussa sekä osa-aikaisena ekonomistina Suomen Pankissa. VTT Henri Nyberg (henri.nyberg@utu.fi) toimii yliopistonlehtorina Turun yliopistossa matematiikan ja tilastotieteen laitoksella. VTT, KTM Harri Pönkä (harri.ponka@vm.fi) toimii erityisasiantuntijana valtiovarainministeriön kansantalousosastolla ja VTT Paolo Fornaro (paolo.fornaro@etla.fi) tutkijana Etlassa. KTM Wilma Nissilä (wilma.v.nissila@utu.fi) toimii taloustieteen tohtorikoulutettavana Turun yliopistossa ja artikkelin kirjoittamisen aikaan myös Suomen Pankissa. VTM Jetro Anttonen (jetro.anttonen@helsinki.fi) toimii tohtorikoulutettavana Helsingin yliopistossa ja tutkimusharjoittelijana Etlassa. Artikkelin perustuu osaltaan Turun yliopistossa (Tilastotieteen keskuksessa) 6.8.2019 pidettyyn seminaariin ”Nowcasting methods and Finnish Business Cycles” ja sen yhteydessä saatuihin kommentteihin. Nyberg kiittää Suomen Akatemiaa (projekti 321968), Emil Aalosen Säätiötä ja OP Ryhmän tutkimussäätiötä taloudellisesta tuesta. Kiitämme Antti Suvantoa ja kahta anonyymiä lausunnonantajaa hyödyllisistä kommentteista. Artikkelissa esitettävät näkemykset eivät välttämättä edusta Suomen Pankin, valtiovarainministeriön tai Etlan kantoja.

kojen lisäksi keskeisenä haasteena on monenlaisen informaation yhdistäminen kokonaiskuvan muodostamiseksi.

Keskeinen teema nowcasting-kirjallisuudessa on ollut suuren muuttujajoukon hyödyntäminen. Yksittäiset aikasarjat sisältävät muuttujakohtaista satunnaisvaihtelua, mikä keskiarvoistuu pois, kun hyödynnetään suurta muuttujajoukkoa. Suurta muuttujajoukkoa hyödyntävien ekonometristen mallien hankaluus on parametrien määrän kasvaminen. Tällöin mallit kyllä selittävät hyvin aineistoa, jolle ne on estimoitu, mutta niiden ennustekyky on useimmiten heikko.

Kirjallisuudessa on esitetty erilaisia versioita dynaamisiin faktorimalleihin luettavista tilastollisista malleista, jotka ratkaisevat ongelman tiivistämällä suuren muuttujajoukon informaation yhteen tai kahteen faktoriin (Forni ym. 2000). Ensimmäinen faktori, joka selittää suurimman osan makrotalouden aikasarjojen vaihtelusta, voidaan monesti tulkita talouden suhdanneindeksiksi. Suhdanneindeksi kuvaa esimerkiksi bruttokansantuotetta paremmin talouden tilaa, sillä se sisältää informaatiota laajemmasta muuttujajoukosta. Tämän artikkelin jaksossa 1 esitellään viime vuosien aikana Helsingin ja Turun yliopistoissa kehitettyä Suomen talouden suhdanneindeksiä, joka perustuu Marianon ja Murasawan (2003) esittämään dynaamiseen faktorimalliin.

Dynaamisessa faktorimallissa suuren muuttuja- ja parametrimäärän ongelma ratkaistaan tiivistämällä informaatiota. Toinen tapa ratkaista ongelma on hyödyntää *shrinkage*-tekniikoita. Makrotalouden analyysissä ja ennustamisessa on jo pitkään hyödynnetty vektoriautoregressiivisiä malleja, joiden estimoinnissa hyödynnetään bayesilaista tilastotiedettä ja erityisesti ns. Litterman (1979) -prioria ja sen

variaatioita. Yksinkertaisimmillaan asetetaan prioria, jossa mallin aikasarjojen oletetaan noudata joko satunnaiskulkuprosesseja tai olevan valkoista kohinaa. Vetämällä moniparametrisen aikasarjamallin parametreja kohti yksinkertaisemman aikasarjamallin parametriervoja pienennetään mallin estimointiepävarmuutta ja liiallista sovitusta estimointiotokseen (*data overfitting*). Näin mallin ennustekyky paranee estimointiotoksen ulkopuolella.

Suomen Pankin ja Etlan nowcasting-mallit perustuvat bayesilaisiin vektoriautoregressiivisiin malleihin. Niitä esitellään artikkelin jaksossa 2 ja 4. Litterman-prioria ja sen variaatioita kehittyneemmän tavan käsitellä suuria tilastoaineistoja antavat koneoppimisen menetelmät. Näitä hyödynnetään Tilastokeskuksen uuden koko kansantalouden pikaestimaatin tuottamisessa (jakso 3).

Nowcasting-menetelmillä ennustetaan muuttujien arvoa tai laajemman suhdanneindeksin samanaikaista tai aivan lähitulevaisuuden kehitystä. Talouspolitiikan kannalta on myös hyödyllistä ymmärtää, missä kohtaa suhdannesykliä ollaan ja kuinka todennäköisesti suhdannetilanne on muuttumassa. Tähän antavat vastauksen osaltaan binäärivastemallit, joissa ennustettavana muuttujana on talouden suhdannetilanne. Binäärivastemallit yleistyivät makroekonometrisessa tutkimuksessa Estrellan ja Hardouvelisin (1991) Yhdysvaltojen taantumajaksoja koskevan tutkimuksen myötä. Binäärivastemallien estimoinnissa on välttämätöntä aluksi määrätä suhdannetilanteen määrittävät talouden suhdannekäänepisteet. Näiden käänneepisteiden muodostamisessa voidaan osaltaan hyödyntää erilaisia nowcasting-menetelmiä sekä asiantuntijatietoon perustuvaa harkintaa. Valtiovarainministeriössä ja Suomen Pankissa on viime aikoina tutkittu binäärivastemallien käyt-

tökelpoisuutta ja suhdannekäännepisteiden asettamista Suomen talouden seurannassa ja alustavia tuloksia esitellään jaksossa 5.

1. Suomen talouden suhdanneindeksi

Viimeisen reilun vuosikymmenen aikana muuttamat dynaamisiin faktorimalleihin perustuvat suhdanneindeksit ovat vakiinnuttaneen asemansa sekä akateemisessa että käytännön reaaliaikaisessa talouden analyysissä. Esimerkiksi Borağan Aruoban, Francis Dieboldin ja Chiara Scottin kehittämä ADS-indeksi on päiväkohmainen Yhdysvaltojen taloudellista aktiiviteettia kuvaava indeksi (Aruoba ym. 2009). Eurocoin-indeksi taas on kuukausikohtainen arvio euroalueen bruttokansantuotteesta. Sen muodostamisesta kuukausittain vastaa Italian keskuspankki (Altissimo ym. 2010).¹

Alun perin Helsingin yliopistossa vuonna 2009 esitetty indeksi oli tiettävästi ensimmäisiä Suomen aineistolla estimoituja dynaamisia faktorimalleja kansantalouden samanaikaista kokonaistaloudellista aktiiviteettia mitattaessa (Lanne ja Nyberg 2009). Tämä kehitystyö tapahtui osana silloista ekonometrisessa kirjallisuudessa meneillään ollutta nopeaa nowcasting-menetelmien kehitystä. Kyseinen indeksi yhdisti neljännesvuosittain julkaistavan reaalisen bruttokansantuotteen ja keskeisten kuukausittain havaittavien makrotaloudellisten muuttujien sisältämän informaation kuukausikohtaiseksi Suomen talouden tilaa kuvaavaksi suhdanneindeksiksi ja samalla ei-havait-

tavan kuukausittaisen bruttokansantuotteen kasvun arvioksi. Indeksi perustui Marianon ja Murasawan (2003) esittämään dynaamiseen faktorimalliin. Heidän, ja monien muiden Yhdysvalloille rakennettujen samanaikaista informaatiota sisältävien suhdanneindeksien tapaan, indeksi rakentui reaalisen bruttokansantuotteen (ainoa neljännesvuosikohtainen muuttuja) sekä kuukausittain havaittavien teollisuustuotantoa ja työllisyyttä kuvaavien muuttujien varaan. Lisäksi Suomen taloudelle keskeisen ulkomaankaupan vuoksi vienti ja tuonti olivat mukana kuukausikohtaisina muuttujina. Kuukausikohtaisten muuttujien käyttämisen vuoksi indeksin arvo voitiin laskea kuukausittain tarvittavien tilastotietojen tullessa saataville. Käytännössä tämä merkitsi noin 1–1,5 kuukauden viivettä reaaliaikaisessa saatavuudessa, mitä voidaan edelleen pitää varsin kilpailukykyisenä BKT:n noin kahden kuukauden julkaisuviiveeseen nähden.

Indeksiä päivitettiin (melko) säännöllisesti kuukausittain aina vuoden 2017 elokuuhun asti. Vuosien 2009–2014 välisen ajan aikana havaittua indeksin toimintaa ja ominaisuuksia käsiteltiin tarkemmin vuonna 2015 julkaistussa artikkelissa (Lanne ja Nyberg 2015).

Kyseisen indeksin päivityksenä Turun yliopistossa on Antti Fredrikssonin, Henri Nybergin ja Jani Sainion toimesta kehitetty päivitettyä Suomen suhdanneindeksiä, jonka lähtökohtaiset perusteet ja käytettävät muuttujat ovat pitkälti samat kuin edeltävässä indeksissä (Fredriksson ym. 2019). Indeksi perustuu edelleen Marianon ja Murasawan (2003) malliin. Aiempaan indeksiin kuuluneiden muuttujien osalta työllisyyden sijaan käytetään nyt työttömien suhteellista osuutta työvoimasta. Muuttujajoukkoa on lisäksi täydennetty rahoitusmarkkinoilta saatavissa olevalla informaatiolla

¹ Ks. tarkemmin <http://www.philadelphiafed.org/research-and-data/real-time-center/business-conditions-index/> ja <http://eurocoin.cepr.org/>.

menneiden osakemarkkinatuottojen muodossa. Osakemarkkinainformaatiota pidetään yleisesti suhdannetilaa ennakoivana indikaattorina, joten tässä tapauksessa osaketuottoja viivästetään siten, että ne kuvaavat paremmin tämänhetkistä talouden tilaa. Rahoitusmarkkinainformaation käyttäminen tarkoittaa käytännössä sitä, että indeksi on mahdollista päivittää minä ajanhetkenä tahansa käytettävän dynaamisen faktorimallin rakenteen vuoksi, mikä vastaa toiminnaltaan esimerkiksi Aruoban ym. (2009) Yhdysvaltojen taloutta koskevaa päiväkohtaista indeksiiä.

Kuviossa 1 esitetään päivitetyn suhdanneindeksin aikaura vuoden 1990 tammikuusta vuoden 2019 toukokuuhun asti. Suomen talouden pitkäaikainen kasvu näkyy indeksin tason nousevana trendinä. Taantuma-ajankohtina indeksin taso laskee, mikä merkitsee siis kokonaisaktiiviteetin supistumista ja siten indeksin negatiivisia kuukausimuutoksia. Näinä kuu-

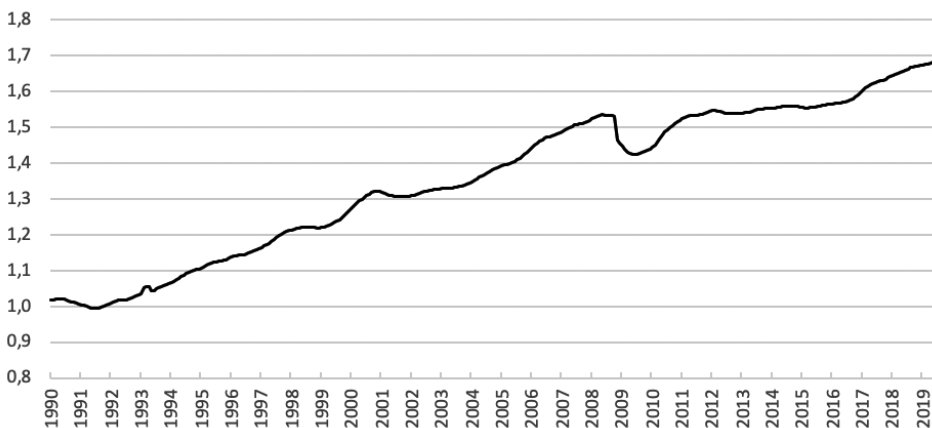
kausina taloudellinen aktiiviteetti on siis supistunut. Merkittävimmät taantumajaksot ajoittuvat odotetusti 1990-luvun alkuun ja vuosiin 2008 ja 2009.

Indeksin nousevasta trendistä johtuen (vuotuiset) kuukausimuutokset (logaritmiset differenssit) ovat tarkasteltavalla ajanjaksolla olleet pääosin positiivisia (keskiarvo 1,72 % ja keskihajonta 3,89). Tämä vastaa varsin tarkoin kansantalouden neljännesvuositilinpitoon perustuvan reaalisen bruttokansantuotteen vuosikasvua (1,59 %).

2. Suomen bruttokansantuotteen lyhyen aikavälin ennustaminen BVAR-mallilla

Suomen Pankin nowcasting-malli ennustaa Suomen reaalisen bruttokansantuotteen kasvua edeltävälle, kuluvalle ja tulevalle vuosineljän-

Kuvio 1. Suhdanneindeksin aikasarja (taso) 1990 (tammikuu) – 2019 (toukokuu)



nekselle.² Malli hyödyntää laajaa, lähes viidenkymmenen muuttujan joukkoa, joka sisältää luottamusindikaattoreita, työmarkkinatilastoja sekä volyyymi- ja liikevaihtoindeksejä.

Mallin tuottaman ennusteen lisäksi mallia hyödynnetään tilastojulkistusten seurannassa. Malli antaa ennusteen kaikille mallin sisältämille muuttujille. Vertaamalla muuttujan ennustetta uuteen tilastojulkistukseen saadaan käsitys, kuinka paljon uutta informaatiota tilastojulkistus sisältää. Mitä enemmän tilastojulkistus eroaa ennustetusta, sitä enemmän bruttokansantuotteen ennuste muuttuu. Tähän vaikuttaa myös muuttujan ja bruttokansantuotteen ennustevirheiden korrelaatio. Mallin tuottama uutisanalyysi auttaa hahmottamaan, mitkä tilastojulkistukset ovat oleellisimpia talouden seurannan kannalta.

Suomen Pankin nowcasting-malli on bayesilainen vektoriautoregressiomalli (BVAR).³ Vektoriautoregressiossa mallin muuttujat ovat endogeenisia ja kunkin muuttujan kuluva hetken arvoja selitetään vakiotermin lisäksi muuttujan omilla ja mallin muiden muuttujien viiveillä. Tästä rakenteesta johtuen VAR-mallin parametrien määrä on yleensä hyvin suuri estimoinnissa käytettävien aikasarjojen pituuden nähden. Tämän vuoksi vektoriautoregressioiden estimoinnissa on tyypillistä hyödyntää bayesilaista tilastotiedettä ja informatiivisia prioreita.

Litterman-priorissa VAR-mallin aikasarjat noudattavat satunnaiskulkuprosesseja, kun kyseessä on tasomuuttujia ja logaritmisoituja muuttujia, kuten Suomen Pankin nowcasting-

mallissa (Litterman 1979). Kun mallin aikasarjat ovat kasvuasteita tai muutoksia, aikasarjojen havainnot generoivien prosessien priorioletus on valkoinen kohina. Satunnaiskulkuprosessin tapauksessa muuttujan oman ensimmäisen viiveen kerroin saa arvon yksi ja muut kertoimet ovat nolliä. Valkoisen kohinan tapauksessa kaikki kertoimet ovat nolliä. Parametrien posteriori-estimaatit ovat priorijakaumien ja aineiston informaation, VAR-mallin tapauksessa pienimmän neliösumman estimaattien, painotettuja keskiarvoja, missä painot riippuvat priorinformaation ja aineiston informaation tarkkuudesta.

Varhaisemmassa kirjallisuudessa tavoitteena oli noin kymmenen muuttujan tai pienempien VAR-mallien tehokas estimointi ja hyödyntäminen ennustamisessa (Litterman 1984; Doan ym. 1984). Tuoreemmassa kirjallisuudessa on näytetty, että vastaavilla menetelmillä voidaan estimoida jopa sadan muuttujan vektoriautoregressiota ja hyödyntää mallia onnistuneesti ennustamisessa (Banbura ym. 2010). Kun moniparametrisen aikasarjamallin parametreja vedetään kohti yksinkertaisemman aikasarjamallin parametriarvoja – kohti ykkösiä ja nolliä – pienennetään mallin estimointiepävarmuutta ja liiallista sovitusta estimointiotokseen (*data overfitting*), ja näin mallin ennustekyky paranee estimointiotoksen ulkopuolella.

Suurin osa Suomen Pankin nowcasting-mallin muuttujista on kuukausitason muuttujia ja osa on neljännesvuositason muuttujia. Tämän vuoksi estimoidaan kaksi VAR-mallia. Ensin estimoidaan ainoastaan kuukausitason muuttujia sisältävä malli, jolla ennustetaan sen neljänneksen loppuun, josta on yhtään kuukausihavaintoa. Kuukausimuuttujat aikaaggregoidaan neljännesvuosimuuttujiksi. Aika-

² <https://www.suomenpankki.fi/nowcast>

³ Mallin tarkempi kuvaus löytyy julkaisusta Itkonen ja Juvonen (2017).

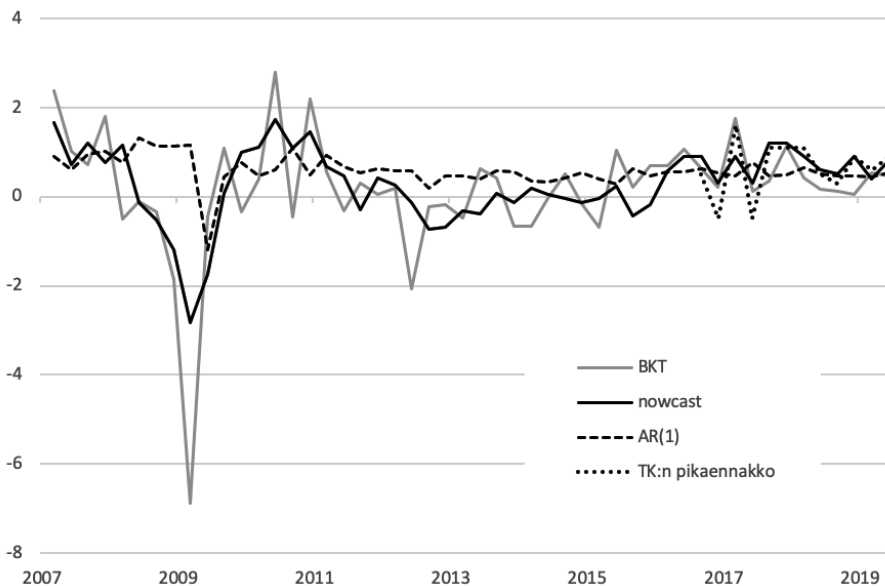
aggregointi tapahtuu Kalman-suotimen avulla, jotta kuukausimuuttujien ennusteita sisältäviin aika-aggregoituihin neljännesvuosimuuttujiin liittyvä ennuste-epävarmuus kyetään ottamaan huomioon. Neljännesvuosimuuttujille estimoidaan VAR, ja Kalman-suodinta hyödyntämällä täytetään puuttuvat havainnot sekä ennustetaan yksi neljännes yli sen neljänneksen, jolta tuoreimmat tilastojulkistukset ovat.

Kuviossa 2 näytetään nowcasting-mallin takautuvasti tehdyt ennusteet vuoden 2007 alusta lähtien vuoden 2016 toiseen neljänneeseen asti sekä reaaliaikaiset ennusteet vuoden 2016 kolmannesta neljänneksestä lähtien vuoden 2019 toiseen neljänneeseen asti. Mallin estimointiotos alkaa vuoden 1995 alusta. Takautuvasti tuotetut ennusteet hyödyntävät uusimpia tilastoja eikä ennustehetken tilastoja.

Ennusteet on tuotettu siten, että ennusteajankohta vastaa ennustettavan vuosineljänneksen loppua, jolloin vuosineljänneksen kaikilta kuukausilta on tiedossa luottamusindikaattorit, mutta muiden muuttujien osalta on puuttuvia havaintoja.

Verrattuna yksinkertaiseen aikasarjamalliin, ensimmäisen asteen autoregressiiviseen prosessiin (AR(1)), BVAR-mallin ennusteet reagoivat nopeammin vuoden 2008 alusta alkavaan bruttokansantuotteen supistumiseen. Vuosien 2011–2014 nollakasvun aikana BVAR-mallin ennusteet ovat keskimäärin lähellä toteutunutta BKT:n kasvua, mutta AR(1)-mallin ennusteet ovat huomattavasti korkeammat. Tarkempaa ennustevirheanalyysia on tehty julkaisussa Itkonen ja Juvonen (2017), missä ennustevirheanalyysissa havaittiin, että hyvin lyhyillä ennuste-

Kuvio 2. Takautuvasti tuotetut ennusteet 2007N1–2016N2 ja reaaliaikaiset ennusteet 2016N3–2019N2



horisonteilla BVAR-mallin ennusteet ovat selkeästi yksinkertaisten aikasarjamallien ennusteita tarkempia, mutta ennustettaessa kuluva vuosineljänneistä kauemmas ero pienenee (Itkonen ja Juvonen 2017). Tulos on yhteneväinen tutkimuskirjallisuuden kanssa.

BVAR-mallilla tuotettuja reaaliaikaisia ennusteita on tallennettu vuoden 2016 kolmannesta neljänneksestä alkaen. Kuviossa 2 ennusteita verrataan Tilastokeskuksen BKT:n pikaennakkoon, joka julkaistaan 45 päivää tilastoitavan vuosineljänneksen päättymisen jälkeen (ks. jakso 3). BVAR-mallin ennusteet ovat pikaennakon julkaisuajankohdan mukaisia. Yllättävästi pikaennakon ja BVAR-mallin ennusteet kulkevat hyvin lähekkäin, vaikka ne tuotetaan hyvin eri tavoin. Vuosina 2016 ja 2017 BVAR-mallin ennusteissa ei esiinny systemaattista harhaa. Vuonna 2018 sen sijaan tämänhetkisten tilastojen perusteella BVAR:n ennusteet kuten myös pikaennakon luvut olivat huomattavasti toteutunutta bruttokansantuotteen kehitystä suurempia.

Suomen Pankin nowcasting-mallin toiminta on kokonaan automatisoitu. Suomen Pankin palvelimella oleva koodipaketti lataa uusimmat tilastot Suomen Pankin aikasarjatietokannasta, kausitasoittaa osan aikasarjoista (aikasarjat, joista ei aineistolähteessä ole saatavilla kausitasoitettua aikasarjaa) sekä suorittaa nowcasting-mallin estimoinnin ja ennusteen tuottamisen. Mallin tuottamat tulokset siirtyvät Suomen Pankin nowcasting-sivustolle sekä Robottiekonomistin twitter-tilille. Tekemällä mallin antamista ennusteista ja uutisanalyysistä julkisia on pyritty havainnollistamaan tilastolliseen malliin perustuen, kuinka paljon yksittäisille uusille tilastojulkistuksille tulisi antaa merkitystä, kun päivitetään käsitystä talouden kehityksestä.

3. Mikroaineistoilla ja koneoppimisen menetelmillä BKT:n pikaestimaatteja

Tilastokeskus ja Etna tekevät yhteistyötä tutkimushankkeessa osana Eurostatin Big Data-konsortiota, jossa yhtenä tavoitteena on luoda uusia tapoja tuottaa nopeampia estimaatteja talousindikaattoreista. Erityisenä tavoitteena on ollut pienentää Tilastokeskuksen julkaiseman tuotannon suhdannekuvaajan julkaisuvivettä 45 päivästä 16 päivään ja näin tuottaa entistä nopeammin myös tuotannon suhdannekuvaajan kehitykseen perustuva ennakkotieto bruttokansantuotteen neljännesvuosikasvusta. Koko talouden pikaestimaatti -nimellä oleva tilasto julkaistaan tällä hetkellä kokeellisena tilastona Tilastokeskuksen sivuilla.⁴

Koko talouden pikaestimaatti tuotetaan hyödyntämällä isoa joukkoa tilastollisia malleja sekä käyttämällä koneoppimisen menetelmiä, jotka sopivat laajojen tilastoaineistojen käsittelyyn.⁵ Pikaestimaatin laskennassa hyödynnetään yritystason aineistoja. Suurin osa muuttujista perustuu Tilastokeskuksen kuukausittain keräämään myyntitiedustelun aineistoon, jossa on mukana noin 2 000 teollisuuden, rakentamisen, kaupan ja palveluiden alojen merkittävintä yritystä. Näiden yritysten liikevaihto kattaa kaiken kaikkiaan noin 70

⁴ <https://www.stat.fi/tup/kokeelliset-tilastot/subdanneltilastot/index.html>

⁵ Kirjallisuudessa käytettyjä menetelmiä ovat mm. erilaiset faktorimallit jakson 1 tapaan (ks. myös Stock ja Watson 2002) sekä ns. shrinkage-menetelmät (kutistamismenetelmät, kuten ridge regression, LASSO ja elastic net) ja muut koneoppimisen algoritmit (boosting, regression trees ja random forests). Tarkemmat esittelyt näistä menetelmistä löytyvät Hastien, Tibshiranin ja Friedmanin kirjasta (2009).

prosenttia kaikkien yritysten liikevaihdosta, mikä tekee aineistosta erinomaisen tietolähteen Suomen talouden suhdannetilän arvioimiseksi. Aineisto on hyödynnettävissä lyhyellä viipeellä, sillä yritykset toimittavat tiedot Tilastokeskukselle viimeistään 16 päivää tilastoitavan kuukauden päättymisen jälkeen.

Estimointi tapahtuu kahdessa vaiheessa. Ensiksi estimoidaan suuri joukko (noin 130) tilastollisia malleja. Näillä malleilla ennustetaan tuotannon suhdannekuvaajan suhteellista muutosta vuodentakaiseen vastaavaan kuukauteen. Ennustekykyä arvioidaan suhteessa toteutuneeseen tuotannon suhdannekuvaajaan. Malleista valitaan noin kaksikymmentä parasta, jotka tuottavat pienimmän ennusteharhan, eli ne eivät systemaattisesti yli- tai aliarvioi tuotannon suhdannekuvaajan kehitystä. Parhaimpien mallien ennusteiden keskiarvosta saadaan tuotannon suhdannekuvaajan pikaestimaatti. Bruttokansantuotteen neljännesvuosimuutoksen pikaestimaatti muodostetaan hyödyntämällä kyseisen neljänneksen kahden ensimmäisen kuukauden osalta tuotannon suhdannekuvaajan tietoja ja viimeisen kuukauden osalta käytettävissä olevaa tuotannon suhdannekuvaajan pikaestimaattia.

Menetelmää on testattu takautuvasti maaliskuusta 2012 joulukuuhun 2018 asti kattavalle ajanjaksolle. Ajanjakson lyhyys johtuu siitä, että pidemmälle ajanjaksolle ei ole saatavilla alkuperäisiä reaaliaikaisia tilastotietoja. Alkuperäisten tilastojen käyttäminen on olennaista menetelmän luotettavuuden arvioimiseksi. Ajanjakson pituus on riittävä ennustekyvyn luotettavaan arviointiin. Tämä siitähän huolimatta, että kyseiseltä ajanjaksolta jää pois kiinnostavia ajanjaksoja, kuten finanssikriisiin liittyvä vuosi 2009, jolloin bruttokansantuotteen taso Suomessa romahti. Ennustevirheana-

lyyssissä pikaestimaatteja verrataan tuotannon suhdannekuvaajan (julkaisuviive 45 päivää) ja neljännesvuositilinpitoon perustuvan bruttokansantuotteen (julkaisuviive 60 päivää) ensimmäisiin tilastojulkistuksiin. Analyysissä ei käytetä tilastojen uusimpia julkistuksia, sillä revisiot voivat johtua muista syistä kuin tilastoaineistojen täydentymisestä.⁶

Taulukossa 1 raportoidaan ennusvirheanalyysin tunnuslukuja. Ensimmäisessä sarakkeessa verrataan tuotannon suhdannekuvaajan pikaestimaattia ja tuotannon suhdannekuvaajan ensimmäistä tilastojulkistusta. Toisessa sarakkeessa verrataan neljännesvuosittaisen bruttokansantuotteen pikaestimaattia ja ensimmäistä tilastojulkistusta. Viimeisessä sarakkeessa verrataan tuotannon suhdannekuvaajaan perustuvaa neljännesvuosittaisen bruttokansantuotteen pikaennakkoa ja sen ensimmäistä tilastojulkistusta. Näin voidaan verrata uuden pikaestimaatin tarkkuutta suhteessa käytössä olleeseen pikaennakkoon.

Tulosten perusteella pikaestimaatit ovat harhattomia, sillä keskivirheet ovat pieniä. Pikaestimaatit eivät siis systemaattisesti yli- tai aliarvioi bruttokansantuotteen tai tuotannon suhdannekuvaajan kehitystä. Pikaestimaatin neliöityjen ennustevirheiden keskiarvon neliöjuuri ja ennustevirheiden itseisarvojen keskiarvo ovat lähellä pikaennakon vastaavia lukuja. Tätä voi pitää merkittävä tuloksena, kun huo-

⁶ *Tilastokeskus tarkentaa tuotannon suhdannekuvaajan ja neljännesvuosikohtaisen bruttokansantuotteen tilastoja, kun koko vuoden kansantulotilastot valmistuvat. Esimerkiksi tuotannon suhdannekuvaajan kuukausihavainnot alkuvuodelta voivat muuttua, kun koko vuoden tilastotiedot valmistuvat ja tuotannon suhdannekuvaaja vuositason tasolla täsmäytetään kyseisen vuoden BKT-lukua vastaavaksi.*

mioidaan lähes kuukauden ero julkaisuviiveessä pikaestimaatin ja pikaennakon välillä.

Kuviosta 3 voidaan visuaalisesti tarkastella bruttokansantuotteen pikaestimaattien kykyä ennustaa sen ensimmäistä tilastojulkistusta. Pikaestimaatit ennakoivat hyvin bruttokansantuotteen vuosikasvuasteen alenemisen vuoden 2012 toisen ja vuoden 2013 toisen vuosineljänneksen välillä. Samoin pikaestimaatit ennakoivat

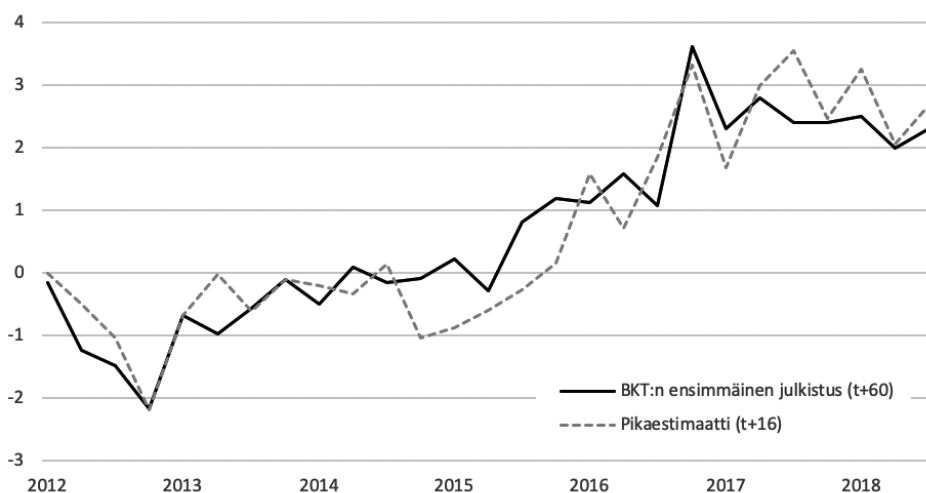
hyvin talouskasvun kiihtymisen vuoden 2016 alussa. Mielenkiintoisesti pikaestimaatit kuitenkin aliarvioivat bruttokansantuotteen kasvua vuoden 2015 aikana, ja tämä ongelma koski myös alkuperäistä bruttokansantuotteen pikaennakkoa kyseisenä vuonna.

Tulokset antavat vaikutelman, että yritystasoaineiston perusteella tuotetut pikaestimaatit ovat suhteellisen laadukkaita, tai ainakin yhtä

Taulukko 1. Ennustevirheanalyysin tunnuslukuja, maaliskuu 2012 – joulukuu 2018

	Pikaestimaatti (kk) vs. tuotannon suhdannekuvaaja	Pikaestimaatti (nv) vs. BKT	Pikaennakko vs. BKT
Keskivirhe	-0.07	0.00	-0.04
Virheiden itseisarvojen keskiarvo	0.86	0.50	0.50
Keskineliövirheen neliöjuuri	1.09	0.63	0.64

Kuvio 3. Bruttokansantuotteen kasvu vuoden takaiseen neljännekseen nähden
Ensimmäinen tilastojulkistus (t+60) ja pikaestimaatti (t+16)



tarkkoja kuin nykyiset Tilastokeskuksen pikaennakot. Pikaestimaattien virheiden itseisarvojen keskiarvo on hyvin lähellä Tilastokeskuksen pikaennakoiden vastaavaa lukua (molemmat noin 0,5 prosenttiyksikköä). Tulokset osoittavat, että epätavanomaisten aineistojen, kuten mikroaineistojen, käyttäminen yhdessä koneoppimismenetelmien kanssa voi olla hyödyllistä, kun tavoitteena on tuottaa nopeampia estimaatteja talouden kehityksestä. Tällä hetkellä Tilastokeskuksessa tutkitaankin tieliikenteen tavarakuljetustilastojen hyödyntämistä pikaennakon tuottamisessa.

4. Google-haut työttömyysasteen ennustamisessa

ETLAnow on Etlassa käynnissä oleva kokeilu hyödyntää uusia aineistolähteitä talouden nykyhetken seurannassa ja lähitulevaisuuden ennustamisessa.⁷ Bruttokansantuotteen sijaan ETLAnow on pääasiassa keskittynyt kausita-soittamattoman työttömyysasteen ennustamiseen. Ennusteet laaditaan Suomen lisäksi koko EU-alueelle. Toinen ETLAnow-projektin puitteissa ennustettu muuttuja on Suomen reaalin asuntohintaindeksi (Widgren 2016).

ETLAnow-projekti sai alkunsa mielenkiinnosta käyttää Google-hakua Suomen työttömyysasteen lähitulevaisuuden ennustamisessa (Tuhkuri 2014). Tarkastelemalla esimerkiksi työttömyysetuuksiin liittyvien hakujen määrän kehitystä, voidaan muodostaa julkaisuviiveen indikaattori työttömyystilanteen kehitykselle. Tämän indikaattorin on havaittu ennustavan työttömyysasteen muutoksia kohtuullisen hyvin etenkin suhdannekäänteissä (Tuhku-

ri 2014). Nykyhetkeä ennustettaessa Google-hakujen merkittävin etu esimerkiksi kyselytutkimuksiin pohjautuviin indikaattoreihin nähden on edellä mainittu julkaisuviiveen puuttuminen. ETLAnow-ennustemalliin Google-hakujen tarjoama informaatio sisällytetään niin sanotun Google-indeksin avulla, joka kaikessa yksinkertaisuudessaan on valittujen hakutermien normalisoitu summa kullakin ajanhetkellä (Tuhkuri 2014; Anttonen 2018).

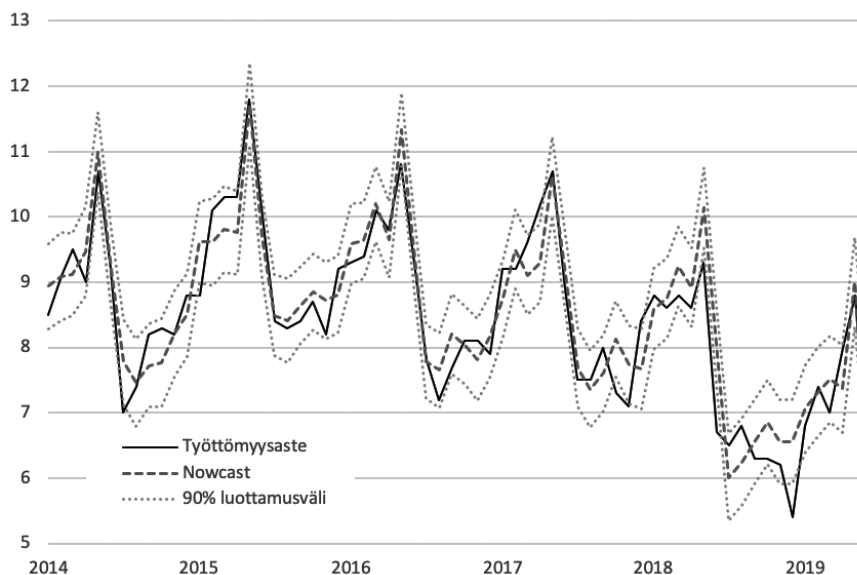
Osana ETLAnow-projektia on tutkittu myös uutisartikkeleiden hyödyntämistä reaaliaikaisten taloudellisten indikaattorien muodostamiseksi. Google-hakujen tapaan uutisartikkelit tarjoavat vaihtoehdoisen julkaisuviiveettömän aineistolähteen talouden nykyhetken tarkasteluun. Uusien aineistojen tarjoamaa lähes reaaliaikaista informaatiota ei kuitenkaan voida hyödyntää tehokkaasti nykyhetken ennustamisessa ilman tehtävään mahdollisimman hyvin sopivia tilastollisia menetelmiä.

ETLAnow-ennustemallin viimeisin versio on bayesiläinen vektoriautoregressio-malli (BVAR), joka käyttää hyväkseen hyvin samankaltaista Litterman-prioria kuin Suomen Pan-kissa nykyhetken ennustamiseen kehitetty bruttokansantuotteen kehitystä seuraava ennustemalli (Itkonen ja Juvonen 2017 ja jakso 2 edellä). Selittäviä muuttujia ETLAnow-mallissa itse työttömyysasteen viipeiden ja Google-indeksin lisäksi ovat kuluttajahintaindeksi ja kyselytutkimuksiin perustuva taloudellista ilmapiiriä mittaava indikaattori (*Economic Sentiment Indicator*, ESI).

Google-aineistoa on saatavilla vasta vuodesta 2004 lähtien, mikä tekee aineistosta kohtuullisen lyhyen ennustemallin estimointiin. Litterman-priorin ansiosta aineiston lyhyys ei ole kuitenkaan ongelma, sillä priorin sisältämällä informaatiolla voidaan kompensoida

⁷ <https://www.etla.fi/etlanow>

Kuvio 4. Toteutunut kausitasoittamaton työttömyysaste ja ETLAnow-mallin ennusteet



aineiston lyhyttä. Estimoinnin lähtökohdaksi ei siis oteta puhtaan epätietoisuuden tilaa, vaan parametrioita, jotka vastaavat hyvin yleisiä makrotaloudellisten aikasarjojen ominaisuuksia, pidetään todennäköisempinä kuin muita parametrioita. ETLAnow-mallin tarkempi tekninen kuvaus on luettavissa projektin verkkosivuilla (Anttonen 2018).

On myös tärkeää muistaa, että ennusteen edustaessa nykytiedon valossa parasta käsitystä tulevaisuudesta (tai nykyisyydestä), ennuste on paljon enemmän kuin usein raportoitu piste-estimaatti. Bayesiläinen tulokulma mallintamiseen mahdollistaakin ETLAnow-mallin tuottamat kokonaiset ennustejakaumat, joita tulisi käyttää päätöksenteossa piste-estimaattien sijaan aina tilanteen niin salliessa.

ETLAnow-mallin ennusteet julkaistaan päivittäin projektin verkkosivuilla. Mallin en-

nustetarkkuuden tarkastelemiseksi myös mallin historialliset ennusteet ovat ladattavissa projektin verkkosivuilla. ETLAnow-mallin ennustetarkkuuden tarkastelu osoittaa mallin tuottaneen useisiin verrokkimalleihin verrattuna huomattavasti tarkempia ennusteita vertailtaessa mallien ennustetarkkuutta vuodesta 2014 nykypäivään (Anttonen 2018). On kuitenkin huomionarvoista, että Googlen hakuaineiston alkaessa vasta vuodesta 2004 ja mallin parametrien määrän ollessa huomattavan suuri, ennustetarkkuuden tarkasteluun käytettyjen kuukausien lukumäärä jää tässä vaiheessa vielä valitettavan pieneksi.

Kuviossa 4 on esitetty ETLAnow-mallin nykyhetken ennusteet ja niiden osumatarkkuus vuodesta 2014 nykypäivään. Ennustetta ympäröivät katkoviivat kuvaavat ennusteen 90 prosentin bayesilaista luottamusväliä (*credible in-*

terval). Nykyhetken ennusteet on laadittu ennusteen kohteena olevan kuukauden viimeisenä päivänä eli noin kolme viikkoa ennen virallisen työttömyysasteen julkaisuajankohtaa. Kuviosta nähdään, että ETLAnow-mallin tuottamat ennusteet työttömyysasteen tasolle ovat huomattavan tarkkoja jo kolme viikkoa ennen virallisen työttömyysasteen julkaisua. Mallin tuottamien ennustetiheyksien luotettavuuden puolesta puhuu myös se, että julkaistu työttömyysaste on osunut mallin tuottamiin 90 prosentin luottamusväleihin noin 90 prosentissa tarkastelluista kuukausista.

5. Suomen talouden suhdannekäänneasteiden määrittäminen ja ennustaminen

Erityisesti finanssikriisin aikaisten ja jälkeisten vuosien voimakkaat taloudellisen aktiviteetin vaihtelut ovat jälleen kasvattaneet kiinnostusta suhdannevaihteluiden mittaamista kohtaan. Suomen taloudelle muodostettujen suhdanetta kuvaavien tilastotietojen, indeksien ja nowcastien perusteella voidaan määrittää suhdannevaiheiden käänneasteet eli tilanteet, joissa suhdannetilä muuttuu vahvasta heikommaksi ja päinvastoin. Paitsi, että ne antavat tiivistetyn kuvan taloustilanteesta tapahtuneista käänneasteista, niiden määrittäminen vastaa osaltaan jatkuvaan keskusteluun taloudellisen taantumamäärityksestä ja tulkinnasta.

Talouspoliittisessa keskustelussa taantumana on perinteisesti pidetty ajanjaksoa, jolloin reaalin bruttokansantuote laskee vähintään kahtena peräkkäisenä vuosineljänneksenä. Sen sijaan erityisesti Yhdysvalloissa on jo pitkään suosittu laajempaa taantumamääritystä, joka perustuu useiden makrotaloudellisten

muuttujien yhteisvaihtelun tarkasteluun. *National Bureau of Economic Research* (NBER) julkaisee Yhdysvaltojen talouden ”viralliset” käänneasteet eli suhdannehuippujen ja suhdannepohjien ajankohdat. Käänneasteiden määrittämisestä vastaava erillinen asiantuntijoista koostuva komitea tarkkailee päätöksenteossaan bruttokansantuotteen ohella useita kuukausittain julkaistavia muuttujia. Taloudellinen toiminta supistuu suhdannehuipun ja -pohjan välisenä ajanjaksona, jolloin talous on taantumassa. Ekspansiovaiheessa, ajanjaksolla suhdannepohjasta suhdannehuippuun, taloudellinen aktiviteetti lisääntyy. Yksi syy laajemman taantumamäärityksen suosimiselle onkin se, että komitea määrittää käänneasteet kuukausitasolla, minkä vuoksi kuukausimuuttujien sisältämän informaation käyttäminen on välttämätöntä.

Käänneasteiden määrittäminen on herättänyt yhä enemmän kiinnostusta viime vuosina. NBER:n ohella muun muassa *Economic Cycle Research Institute* (ECRI) julkaisee käänneasteet 21 maalle. Sen lisäksi *Centre for Economic Policy Research* (CEPR) on perustanut erillisen komitean, joka määrittelee euroalueen suhdannehuiput ja -pohjat. Suomessa ei toistaiseksi ole olemassa ”virallista” käänneastekronologiaa. Tässä artikkelissa esitettyjä suhdanneindeksejä ja nowcastingiin käytettäviä muita lyhyen aikavälin tilastollisia malleja voidaan kuitenkin käyttää Suomen talouden käänneasteiden määrittämiseen, kun on ensin valittu jokin sääntö tai tilastollinen menetelmä, joiden avulla indekseistä tunnistetaan suhdannehuiput ja -pohjat.

Bryn ja Boschanin (1971) kehittämä algoritmi (BB-algoritmi) on yksi käytetyimmistä menetelmistä käänneasteiden määrittämisessä. BB-algoritmi perustuu käänneasteiden etsimi-

Taulukko 2. Kansantalouden suhdannekäännepesteet Suomessa 1988–2018

Otos	Reaalinen BKT		Suhdanneindeksi (Lanne ja Nyberg 2009)		Suhdanneindeksi (Fredriksson ym. 2019)		Tuotannon suhdannekuvaaja	
	1988Q1–2018Q4		1988M1–2017M6		1989M2–2019M3		1995M1–2019M3	
	Huippu	Pohja	Huippu	Pohja	Huippu	Pohja	Huippu	Pohja
1	1990Q1	1993Q2	1990M4	1991M9	1990M5	1991M6		
2					2000M11	2001M8		
3	2007Q4	2009Q2	2008M4	2009M7	2008M5	2009M6	2007M12	2009M5
4	2012Q1	2013Q1	2011M11	2013M3	2012M2	2012M8	2010M12	2014M12
5	2013Q3	2015Q1	2013M12	2015M4	2014M8	2015M4		

seen liukuvan keskiarvon menetelmällä tasoitettua sarjasta. Löydettyjen käännepesteiden on toteutettava useita erilaisia rajoitteita ja päätössääntöjä, jotka liittyvät erityisesti taantumaja ekspansiovaiheiden keston. Ensinnäkin tyypillisesti asetettavat vaatimukset ovat, että taantumajan tai ekspansiojakson (taloudellista kasvua vastaavan ajanjakson) on kestettävä vähintään viisi kuukautta ja suhdannesyklin (aika huipusta huippuun tai pohjasta pohjaan) on puolestaan vähintään 15 kuukautta. Lisäksi käännepesteet voidaan havaita vasta kuuden kuukauden viipeellä algoritmin ollessa kaksisuuntainen (menneisyys- ja tulevaisuusriippuvainen). Harding ja Pagan (2002) ovat kehittäneet BB-algoritmista neljännesvuosiaineistolle soveltuvan version (BBQ-algoritmi)⁸, sillä al-

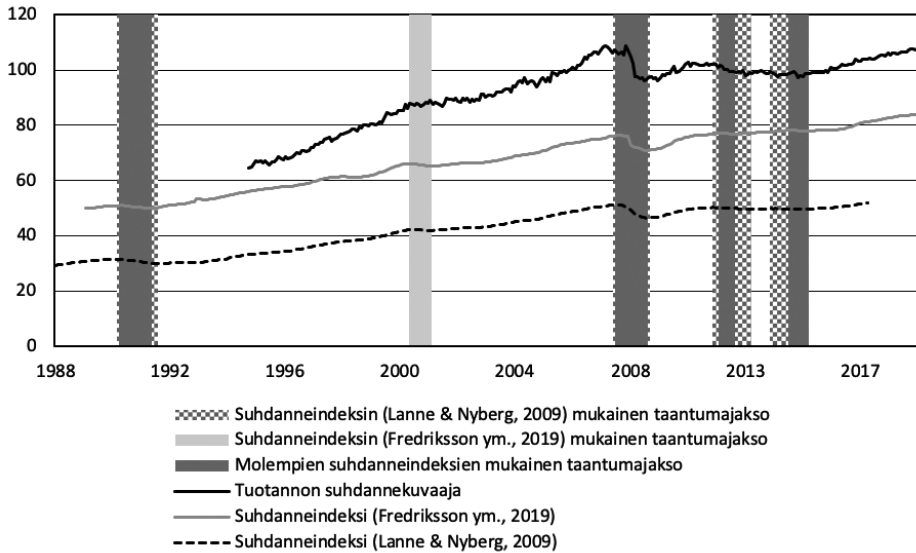
kuperäinen BB-algoritmi on sovellettavissa sellaisenaan vain kuukausittain havaittaville muuttujille.

Taulukossa 2 esitetään BBQ- ja BB-algoritmeilla sekä eri aineistoilla saadut suhdannekäännepesteet Suomen taloudelle ja kuviossa 5 esitetään suhdanneindeksien käännepesteitä vastaavat taantumajaksot. Eri aineistot näyttävät tuottavan varsin odotetut ja samansuuntaiset suhdannehuiput ja -pohjat muutamia poikkeuksia lukuun ottamatta.

Osaketuotoilla täydennetyt suhdanneindeksin mukaan Suomessa koettiin 2000-luvun alussa taloudellisen aktiviteetin laskukausi, mutta muiden muuttujien perusteella varsinaisesta taantumasta ei kuitenkaan ollut kyse. Esimerkiksi reaalin bruttokansantuote ei laskenut kahtena peräkkäisenä neljänneksenä.

⁸ BBQ-algoritmi etsii käännepesteet tasoittamattomasta sarjasta, mutta sisältää muuten kaikki samat vaatimukset ja rajoitteet kuin BB-algoritmikin.

Kuvio 5. Suhdanneindeksien ja BB-algoritmin mukaiset taantumajaksot Suomessa 1988–2018



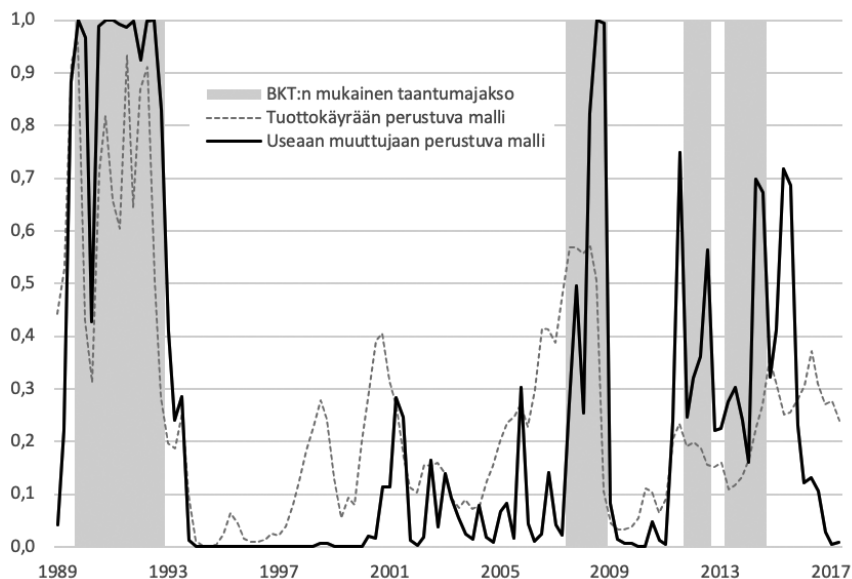
Finanssikiiriin ja sitä seuranneeseen 2010-luvun taantumaa on suomalaisessa talouskeskustelussa viitattu perinteisesti kaksoistaantumana. Suhdanneindeksien ja bruttokansantuotteen perusteella näyttäisi kuitenkin olevan perusteltua puhua kolmoistaantumasta, sillä ainoastaan tuotannon suhdannekuvaajan mukaan 2010-luvun taantumaa voidaan pitää yhtenä pitkäkestoisempaa taantumana.

Aineistot näyttävät ajoittavan myös 1990-luvun suhdannepohjan eri ajanhetkiin. Aikaisemmin Lanne ja Nyberg (2009) ajoittivat suhdanneindeksin ja BB-algoritmin perusteella 1990-luvun taantumaa päätymisen ajankohdaksi maaliskuun 1993, mutta aineiston tarkentumisen (revisioiden) jälkeen indeksi ajoittaa nyt suhdannepohjan syyskuulle 1991. Vastaavasti reaalisuuden bruttokansantuotteen kas-

vuun perustuva kronologia (BBQ-algoritmi) ajoittaa suhdannepohjan vuoden 1993 toiselle neljännekselle, joka kuvastaa sitä, että eri muuttujiin perustuvat käännepesteet voivat poiketa toisistaan selvästi. Tuotannon suhdannekuvaaja on saatavilla vasta vuodesta 1995 alkaen, joten siitä saatavaa informaatiota ei voida hyödyntää 1990-luvun alun taantumaa ajoittamisessa.

Suhdanteiden käännepesteiden perusteella talouden tilasta saadaan muodostettua binäärimuuttuja, jota voidaan hyödyntää muun muassa taantumajaksojen ennustamisessa. Kirjallisuudessa probit- ja logit-malleja onkin käytetty laajasti sen jälkeen, kun Estrella ja Hardouvelis (1991) havaitsivat niiden hyödyllisyyden Yhdysvaltojen taantumajaksoja ennustettaessa. Probit- ja logit-mallien avulla pystytään mallintamaan

Kuvio 6. Probit-mallin tuottamat estimointiotoksen sisäiset (in-sample) taantumataodennäköisyydet 1989–2017



todennäköisyyttä taantumajakson esiintymiselle valittuna ajanhetkenä tai tietyn aikaikkunan sisällä.

Suhdannekäänteiden ennakoimisessa on hyödynnetty useita muuttujia. Estrellan ja Hardouvelisin (1991) tulokset ja heidän työtään seurannut tutkimuskirjallisuus on korostanut erityisesti tuottokäyrän, eli pitkän (esim. 10-vuotisen valtion velkakirjan) ja lyhyen koron (esim. 3 kk) välisen eron, ennustevoimaa. Viime aikoina keskustelu tuottokäyrän ennustekyvystä on ollut vilkasta, kun Yhdysvaltojen tuottokäyrä on kääntynyt negatiiviseksi, minkä johdosta ennustemallien tuottamat taantumataodennäköisyydet ovat nousseet.

Tuottokäyrän lisäksi myös muiden rahoitusmuuttujien, kuten lyhyiden korkojen sekä osakemarkkinoiden tuottojen on havaittu en-

nakoivan suhdannemuutoksia (Estrella ja Mishkin 1998; Nyberg 2010). Muihin yleisesti hyödynnettyihin muuttujiin ovat kuuluneet muun muassa kuluttajien sekä yritysten luottamusta kuvaavat indikaattorit (Christiansen 2014), luottomuuttujat (Pönkä 2017), sekä asuntomarkkinoiden hintojen kehitystä kuvaavat indikaattorit.

Pönkä ja Stenborg (2019) tutkivat hiljattain taantumajaksojen ennustettavuutta Suomessa kirjallisuudessa yleisesti käytettyjen muuttujien pohjalta ajanjaksolla 1988–2017. Tutkimuksessa hyödynnettiin neljännesvuosittaista aineistoa ja suhdannekronologia perustui reaalisen BKT:n kasvuun ja BBQ-algoritmiin. Tulosten perusteella muun muassa tuottokäyrässä ja pääkaupunkiseudun asuntojen reaalihinnoissa tapahtuvat muutokset ennakoivat suhdannekäänteitä.

Yhden muuttujan mallit eivät kuitenkaan kyenneet ennakoimaan kaikkia taantumia, mikä korostaa sitä, ettei täysin luotettavaa yksittäistä, taantumaa ennakoivaa indikaattoria löydy. Parhaisiin ennusteisiin päästään tyypillisesti malleilla, joissa hyödynnetään useita muuttujia⁹.

Kuviossa 6 esitetään sekä tuottokäyrään perustuvan että muita muuttujia hyödyntävän probit-mallin tuottamat taantumadennäköisyydet ja reaalisen bruttokansantuotteen käännepesteitä vastaavat taantumajaksot aikavälillä 1989Q3–2017Q4. Mallit ja niissä käytettävien muuttujien viipeet on esitelty tarkemmin Pöngän ja Stenborgin tutkimuksessa (Pönkä ja Stenborg 2019, taulukko III malli 1, taulukko V malli 33). Tuottokäyrä kääntyi negatiiviseksi ennen Suomen 1990-luvun lamaa sekä ennen kansainvälisen finanssikriisin seurauksena alkanutta taantumaa, minkä vuoksi mallin tuottamat taantumadennäköisyydet nousivat. Sen sijaan 2010-luvulla esiintynyttä kahta taantumaa ei edeltänyt tuottokäyrän kääntyminen negatiiviseksi. Näyttäisikin siltä, että tuottokäyrän ennustevoima (*in-sample*) Suomessa on heikentynyt rahapoliittisten korkojen saavuttaessa nollarajan (*zero lower bound*). Kun malliin lisätään tuottokäyrän ohelle selittäviksi muuttujiksi osaketuotot, pääkaupunkiseudun asuntojen reaali hinnat sekä kuluttajien odotukset työttömyydestä, saadaan huomattavasti parempia sovitteita.

⁹ Suomen Pankissa kesällä 2019 tehdyn, vielä julkaisematoman tutkimuksen mukaan Pöngän ja Stenborgin (2019) tulokset pätevät myös silloin, kun sovelletaan kuukausiaineistoa ja Lanne–Nyberg-subdanneindeksin (2009; 2015) taantumajaksvoja.

6. Lopuksi

Olemme koonneet tähän artikkeliin Suomessa viime aikoina tehtyä tutkimus- ja kehitystyötä talouden tilan mittaamiseen ja lyhyen aikavälin ennustamiseen (*nowcastingiin*) liittyen. Usealla taholla on tehty tavoitteellista kehitystyötä, minkä tuloksena kansainvälisen ekonometrisen tutkimuskirjallisuuden menetelmiä on voitu onnistuneesti soveltaa Suomen talouden seurantaan ja analyysiin.

Suhdanneindeksien perusteella voidaan määritellä kuukauden tai neljännesvuoden tarkkuudella Suomen kansantalouden suhdannevaiheiden käännepesteet käytännön taloudellisen päätöksenteon ja taloustieteellisen tutkimuksen tarpeisiin. Esimerkiksi vuosien 1990–2018 aikana Suomen talous on kokenut useita suhdannekäänteitä, jotka eri menetelmät näyttävät ajoittavan varsin odotetusti. Uusia tilastollisia menetelmiä hyödyntämällä bruttokansantuotteen ennakkotietojen julkaisuviivettä on voitu lyhentää. Bruttokansantuotteen lyhyen aikavälin ennustamisessa pystytään nyt aikaisempaa paremmin hyödyntämään lyhyellä viiveellä julkaistavia tilastoja ja seuraamaan tilastojulkistusten vaikutusta bruttokansantuotteen ennusteen muutokseen. Google-haut työttömyysasteen ennustamisessa parantavat työmarkkinatilanteen reaaliaikaista seuranta.

Laskentakapasiteetin kasvaminen, suurten ja uusien tilastoaineistojen hyödyntäminen sekä menetelmien kehittyminen antavat uusia mahdollisuuksia talouden seurantaan ja ennustamiseen tilastollisin menetelmin. Kun resurssit alan opetus- ja tutkimustyöhön sekä kehitysprojekteihin ovat riittävät, voidaan myös tulevaisuudessa soveltaa uusimpia menetelmiä Suomen talouden seurantaan ja ennustamiseen. □

Kirjallisuus

- Altissimo, F., Cristadoro, R., Forni, M., Lippi, M. ja Veronese, G. (2010), “New Eurocoin: Tracking economic growth in real time”, *Review of Economics and Statistics* 92: 1024–1034.
- Anttonen, J. (2018), “Nowcasting the Unemployment Rate in the EU with Seasonal BVAR and Google Search Data”, ETLA Working Papers 62.
- Aruoba, S. B., Diebold, F. X. ja Scotti, C. (2009), “Real-time measurement of business conditions”, *Journal of Business and Economic Statistics* 27: 417–427.
- Banbura, M., Giannone, D. ja Reichlin, L. (2010), “Large Bayesian vector auto regressions”, *Journal of Applied Econometrics* 25: 71–92.
- Banbura, M., Modugno, M. ja Reichlin, L. (2013), “Now-casting and the real-time data-flow”, teoksessa Elliott, G. ja Timmermann, A. (toim.), *Handbook of Economic Forecasting*, Elsevier-North Holland, 2. painos: 195–237.
- Bry, G. ja Boschan, C. (1971), “Cyclical analysis of time series: selected procedures and computer programs”, National Bureau of Economic Research.
- Christiansen, C., Eriksen J. N. ja Möller, S. T. (2014), “Forecasting U.S. recessions: The role of sentiment,” *Journal of Banking and Finance* 49: 459–468.
- Doan, T., Litterman, R. ja Sims, C. (1984), “Forecasting and conditional projection using realistic prior distributions” *Econometric Reviews* 3:1:100.
- Estrella A. ja Hardouvelis G. A. (1991), “The term structure as a predictor of real economic activity”, *Journal of Finance* 46: 555–576.
- Estrella A. ja Mishkin F. S (1998), “Predicting U.S. recessions: Financial variables as leading indicators”, *Review of Economics and Statistics* 80: 45–61.
- Fredriksson, A., Nyberg, H., ja Sainio, J. (2019), “Coincident Economic Index for the Finnish Economy”, Käskirjoitus, Turun yliopisto.
- Forni, M., Hallin, M., Lippi, M. ja Reichlin, L. (2000) “The Generalized Dynamic-Factor Model: Identification and Estimation”, *The Review of Economics and Statistics* 82: 540–554.
- Ghysels, E., Sinko, A. ja Valkanov, R. (2007), “MIDAS regressions: Further results and new directions”, *Econometric Reviews* 26: 53–90.
- Harding, D. ja Pagan, A. (2002), “Dissecting the cycle: a methodological investigation”, *Journal of Monetary Economics* 49: 365–381.
- Hastie, R., Tibshirani ja Friedman, J. (2009), *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2. painos, Springer.
- Itkonen, J. ja Juvonen, P. (2017), “Nowcasting the Finnish economy with a large Bayesian vector autoregressive model”, *BoF Economics Review* 6/2017.
- Lanne, M. ja Nyberg, H. (2009), “Suomen kansantalouden suhdanneindeksi”, *Kansantaloudellinen aikakauskirja* 105: 422–432.
- Lanne, M. ja Nyberg, H. (2015), “Suomen kansantalouden suhdanneindeksi 2009–2014 ”, *Kansantaloudellinen aikakauskirja* 111: 6–15.
- Litterman, R. (1979), “Techniques of forecasting using vector autoregressions”, Federal Reserve Bank of Minneapolis Working Paper. no. 115.
- Litterman, R. (1984), “Specifying VAR’s for macroeconomic forecasting”, Federal Reserve Bank of Minneapolis Staff Report 92.
- Marcellino, M. (2006), “Leading indicators”, teoksessa Elliott, G., Granger, C. W. J. ja Timmermann, A. (toim.), *Handbook of Economic Forecasting*, Elsevier, 1. painos: 879–960.
- Mariano, R. S. ja Murasawa, Y. (2003), “A new coincident index of business cycles based on monthly and quarterly series”, *Journal of Applied Econometrics* 18: 427–443.

- Nyberg, H. (2010), “Dynamic probit models and financial variables in recession forecasting”, *Journal of Forecasting* 29: 215–230.
- Pönkä, H. (2017), “The role of credit in predicting U.S. recessions”, *Journal of Forecasting*, 36: 221–247.
- Pönkä, H. ja Stenborg, M. (2019), “Forecasting the state of the Finnish business cycle”, Valtiovarainministeriön julkaisuja 2019:13.
- Stock, J. ja Watson, M.W. (2002) “Forecasting using principal components from a large number of predictors”, *Journal of the American Statistical Association* 97: 1167–1179.
- Tuhkuri, J. (2014), “Big Data: Google-haut ennustavat työttömyyttä Suomessa”, ETLA Raportit 31.
- Widgrén, J. (2016), “Google-haut Suomen asuntojen hintojen ennustajana”, ETLA Raportit 63.