

Syntyvyyden lasku ja sen mahdolliset selitykset - LUONNOS

Antti Tanskanen

3.5.2018

Johdanto

Syntyneiden lasten määrä on Suomessa laskenut 61 000 lapsesta vuonna 2010 noin 50 300:ään vuonna 2017. Laskeva trendi on jatkunut koko aikavälin 2010-2017, joten satunnaisvaihtelusta tuskin on kysymys. Osan laskusta selittää synnytysikäisten naisten lukumäärän pieneminen, mutta ei läheskään kaikkea. Muiksi selityksiksi on tarjottu mm. työttömyyttä (Hiilamo, 2017) ja miesten heikkoa työsuheteeseen kiinnitymistä (Miettinen ja Jalovaara, 2015). Muissa maissa syntyvyyden laskun on arvioitu aiheutuvan mm. lapsikuolleisuuden vähenemisestä (Becker ja Barro, 1988), kaupungistumisesta (Martine et al., 2013) ja vanhuuseläkkeiden korkeammasta korvausasteesta (Boldrin et al., 2015). Sen lisäksi että työttömyys alentaisi syntyvyyttä, taloudellisten kriisien on myös havaittu lisänneen syntyvyyttä (Kohler ja Kohler, 2002).

Syntyneiden määrällä on viiveillä vaikutusta työikäisen väestön määrään, ja sitä kautta syntyvyyden laskulla on vaikuttuksia mm. sosiaaliturvan rahoituksen kestävyyteen. Eläketurvakeskuksen laskelmien mukaan työeläkemaksussa on pitkällä aikavälillä useiden prosenttiyksiköiden nousupaine (Tikanmäki, 2017).

Tässä tutkimuksessa tavoitteena on hakea eri tekijöiden merkitystä syntyvyyden laskuun ja tarkastella, onko yksioikoisella regressiolla mahdollista löytää aineistosta viitteitä siitä, että jokin tai jotkin esitetyistä syistä selittäisivät syntyvyyden laskua. Koska aineisto on perin rajallista, on tuloksiin syytä suhtautua varovaisesti.

Aineisto

Tutkimuksessa käytetty aikasarja-aineisto on peräisin tilastokeskuksen StatFin-tietokannasta, ETK:lta ja Human Mortality Databasesta. Aineisto koostuu vuosien 1989-2017 datasta. Väestörakenteesta (Tilastokeskus) mukana on syntyneidet lukumäärät, kokonaishedelmällisyysluvut, synnyttäjien keski-ikä ja naisten lukumäärät 5-vuotisikäryhmittäin. Mukana ovat myös avioliittojen ja avioerojen määrät (Tilastokeskus).

Työllisyyttä (Tilastokeskus) kuvataan 10-vuotisikäuokittaisilla työllisyysasteilla sukupuolittain, 10-vuotisikäuokkakohtaisista työttömyysasteista sukupuolittain, ja työsuhteiden tyypitiedoilla (jatkuva/määrääkainen; kokoikainen/osa-aikainen) vuosilta 1997-2017. Pienituloisuutta kuvaa 10-vuotisikäuokittainen pienituloisten osuus (tulos alle 60 % mediaanitulosta) väestöstä (Tilastokeskus). Kaupungistumista kuvaa maalla, ulommassa kaupungissa ja sisemmässä kaupungissa asuvien osuudet väestöstä (Tilastokeskus). Eläkejärjestelmän vaikutusta syntyvyyteen yritetään selvittää käytämällä vertailukohtana eläkemenoa BKT:stä (Eläketurvakeskus). Lapsikuolleisuus (Human mortality database)

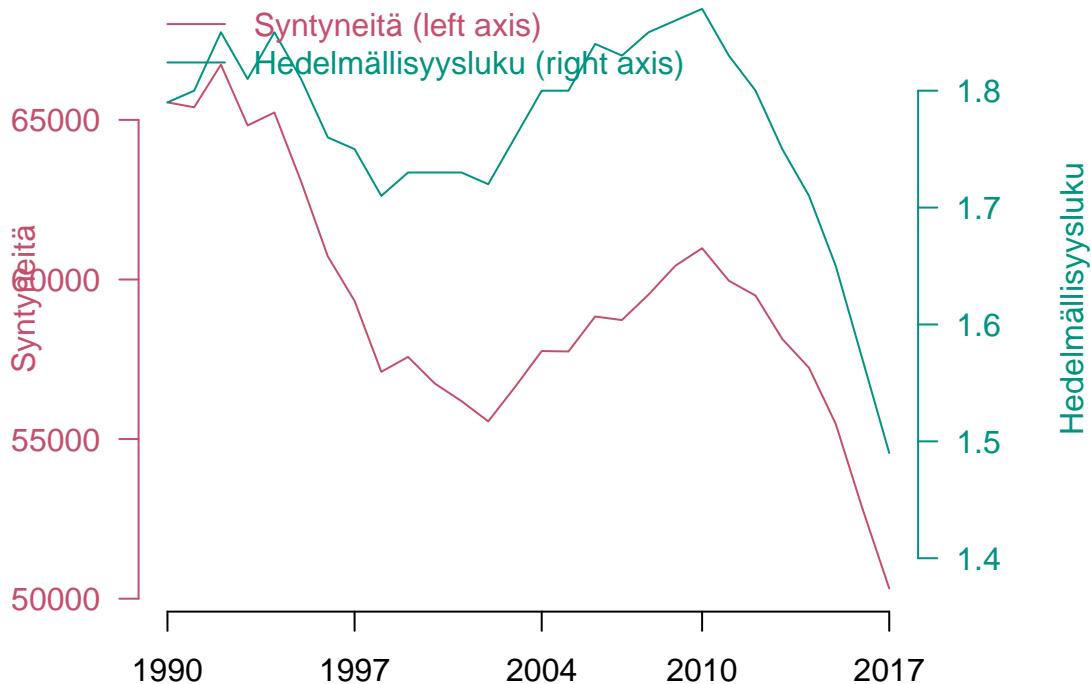
Kaikkiaan lähtöaineistossa on 39 aikasarjaa, joissa on 29 havaintoa vuosilta 1989-2017, tosin osasta aikasarjoja vuosien 1989 ja/tai 2017 luvut puuttuvat. Pidempiä aikasarjoja löytyy syntyneistä, naisten 5-vuotisikäuokittaisesta määristä ja hedelmällisyysluvusta. Näistä käytetään aikasarjoja vuosilta 1970-2017.

Aineiston perusteella ei voi tehdä kausaalisia päätelmiä, ainostaan assosiaatiivisutta koskevia. Tämä tietenkin rajoittaa tulkintamahdollisuksia.

Tarkoituksesta on avoimen, toistettavan tutkimuksen hengessä näyttää täsmällisesti mitä on tehty. Tästä syystä alkuperäinen aineisto (myös tämä työkirja koodeineen, Rmd-muotoisena; data Rdata-muodossa) on kokonaisuudessaan saatavissa osoitteesta <https://github.com/ajtanskanen/Syntyvyyden-tekij-t>

Aikasarjojen vertailut

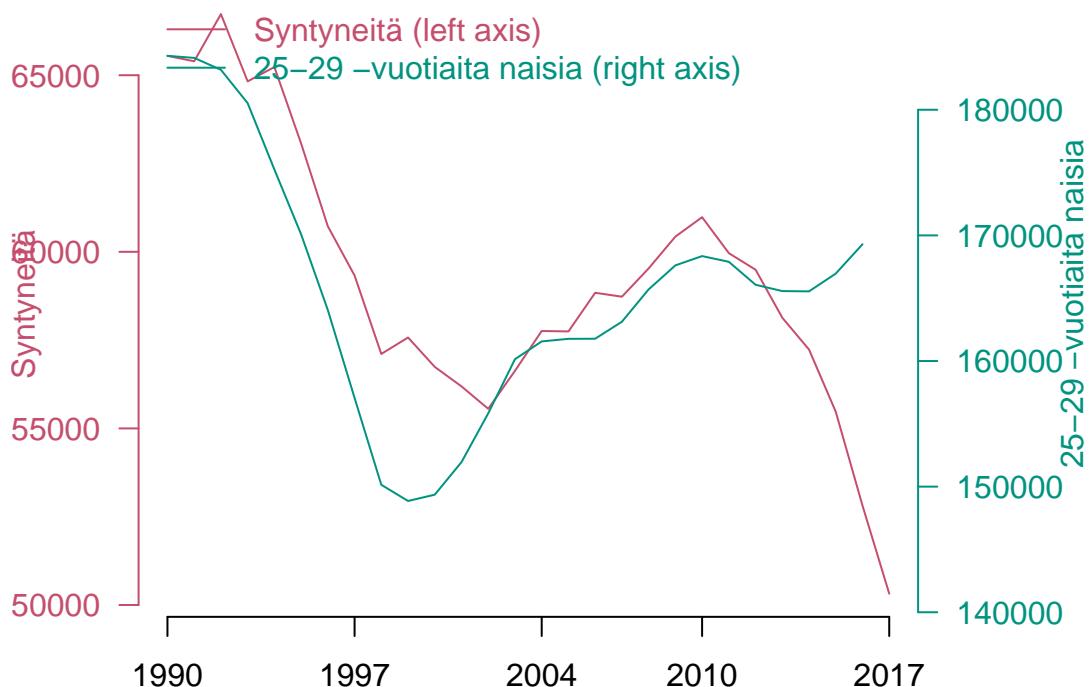
Syntyneiden määrä noudattelee melko hyvin muutoksia hedelmällisyysluvussa



```
## $mar
```

```
## [1] 3.1 6.1 3.1 6.1
```

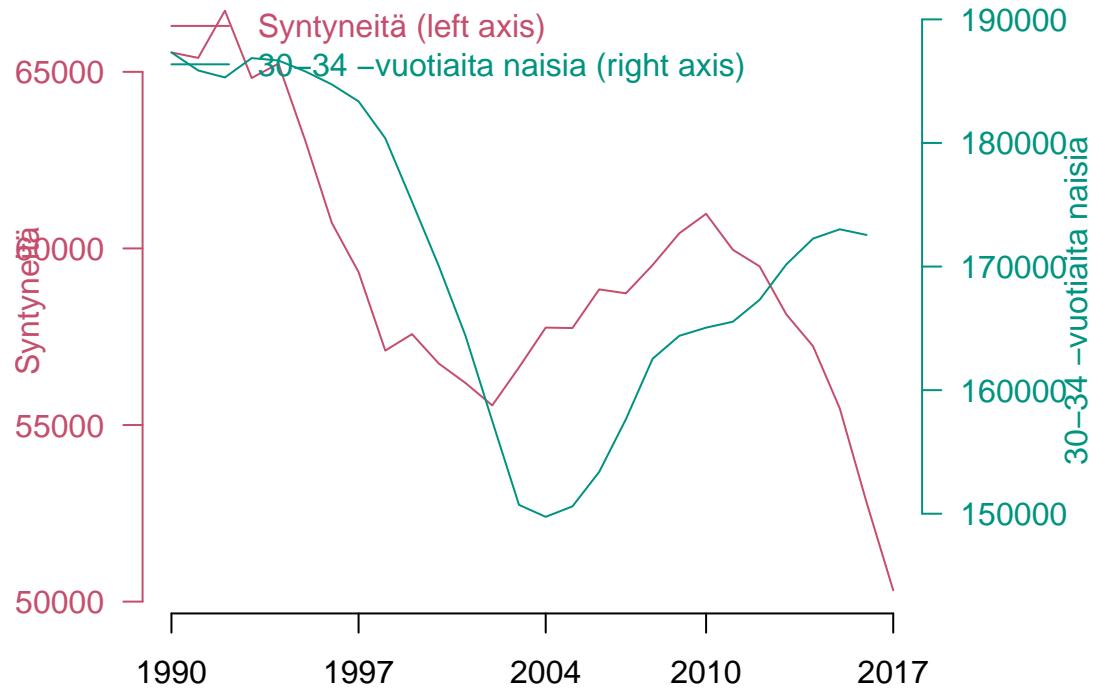
Toinen syntyneiden määrää selittävä tekijä on synnytyskäisten naisten lukumäärä. Kun tarkastellaan 25-29-vuotiaiden naisten määrää, näkyy, että vuoteen 2010 asti köyrät kulkevat yhtä matkaa. Sen jälkeen ne erkautuvat.



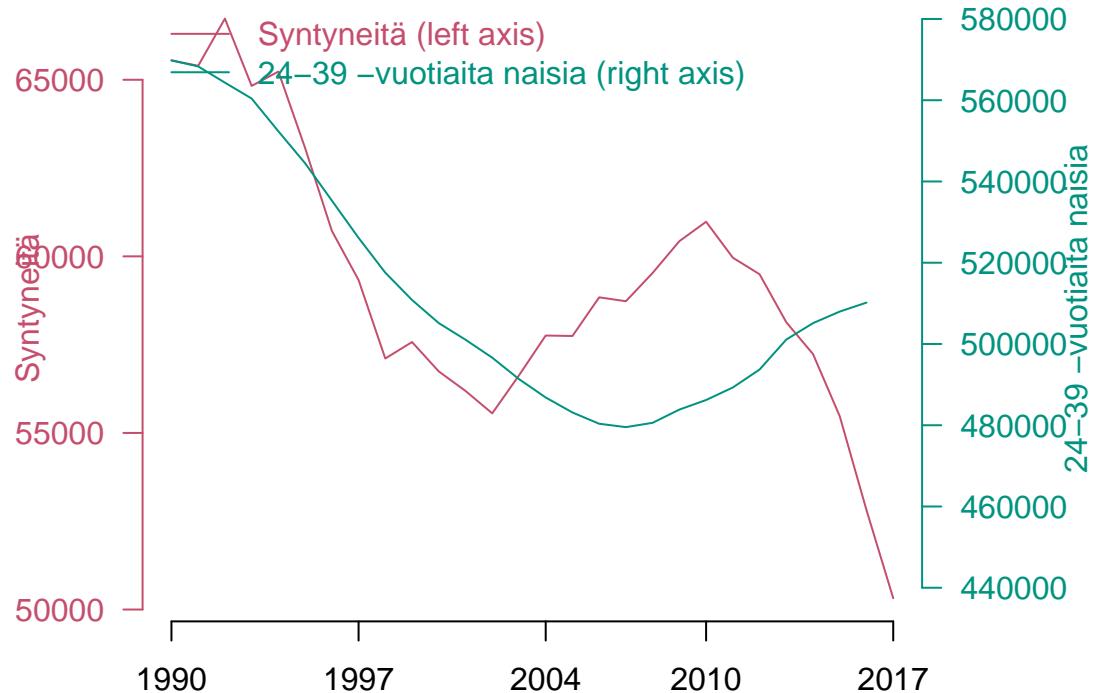
```
## $mar
```

[1] 3.1 6.1 3.1 6.1

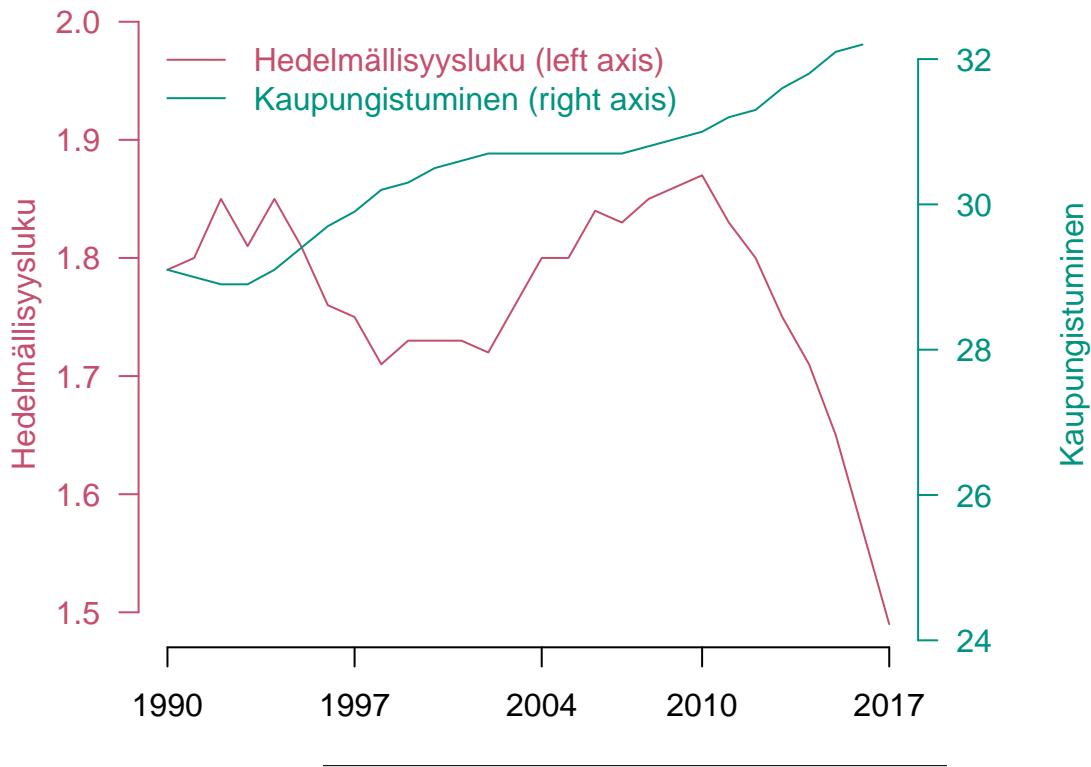
Sama ilmiö näkyy myös 30-34 -vuotiaiden naisten määrässä



Ja myös 35-39 -vuotiaiden naisten määrässä



Urbaanistiasuvien osuus Suomessa on vahvasti negatiivisesti korreloitunut hedelmällisyyslувун kanssa



Mikä selittää vaihtelun syntyneiden lasten määrässä?

Aloitetaan yksinkertaisesta mallista syntyneiden lukumäärälle: mukana vain hedelmällisyysluku ja synnytysikäisten naisten lukumäärät

```
syntyvyys_base = bas.lm(syntyneita ~ hedelmällisyysluku + naisten_lkm_15_19 + naisten_lkm_20_24 +naisten_lkm_25_29 +naisten_lkm_30_34 +naisten_lkm_35_39 +naisten_lkm_40_45 +naisten_lkm_45_49 , data = syntyvyysaineisto_redusoitu, prior = "ZS-null", modelprior = uniform())
```

Sovite on bayesilainen, jossa eri mallien prioritodennäköisyytenä on käytetty samaa arvoa. Sovite antaa todennäköisyydet malleille.

	P(B != 0 Y)	model 1	model 2	model 3	model 4	model 5
## Intercept		1.000	1.000	1.000	1.000	1.000
## hedelmällisyysluku		1.000	1.000	1.000	1.000	1.000
## naisten_lkm_15_19	0.018	0.000	0.000	0.000	1.000	0.000
## naisten_lkm_20_24	0.986	1.000	1.000	1.000	1.000	0.000
## naisten_lkm_25_29	1.000	1.000	1.000	1.000	1.000	1.000
## naisten_lkm_30_34	1.000	1.000	1.000	1.000	1.000	1.000
## naisten_lkm_35_39	1.000	1.000	1.000	1.000	1.000	1.000
## naisten_lkm_40_45	0.028	0.000	1.000	0.000	0.000	0.000
## naisten_lkm_45_49	0.033	0.000	0.000	1.000	0.000	1.000
## BF	NA	1.000	0.025	0.021	0.017	0.009
## PostProbs	NA	0.926	0.023	0.020	0.016	0.008
## R2	NA	0.999	0.999	0.999	0.999	0.998
## dim	NA	6.000	7.000	7.000	7.000	6.000
## logmarg	NA	63.946	60.252	60.094	59.894	59.179

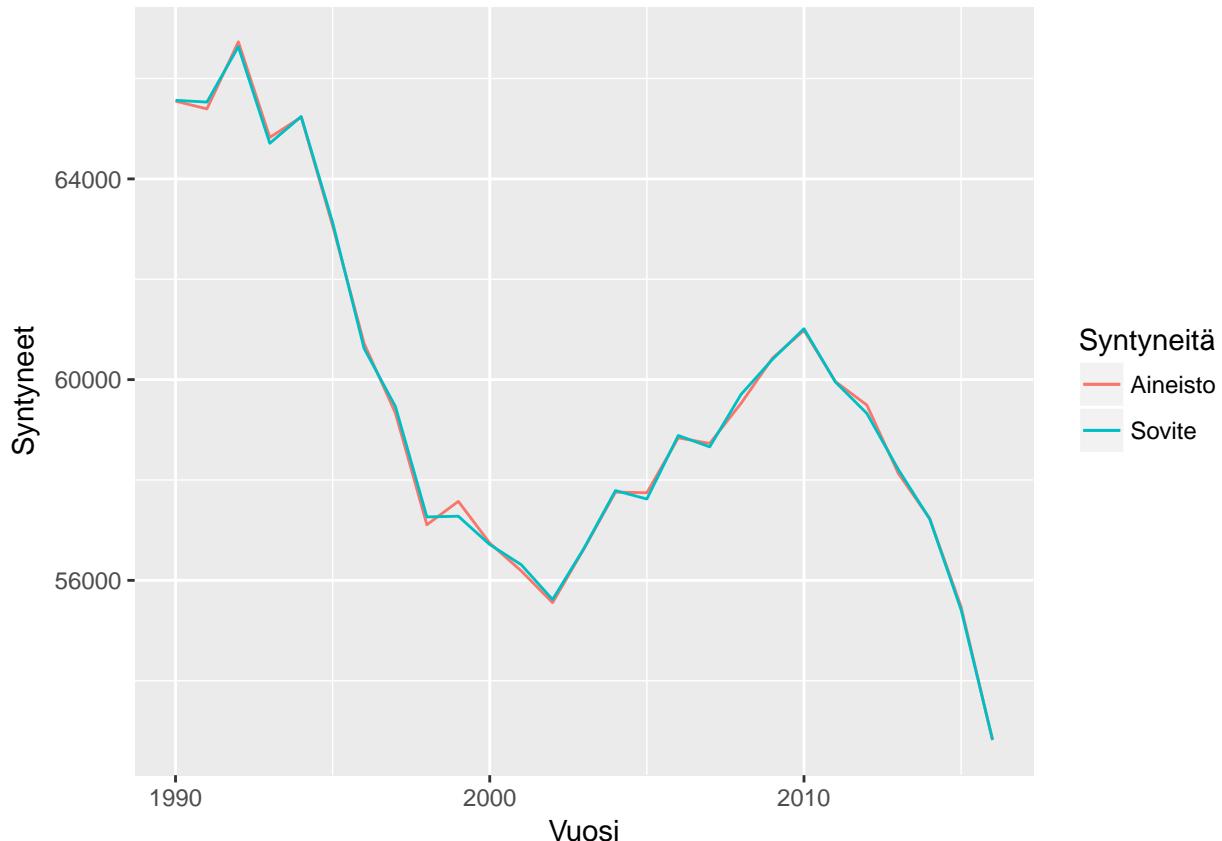
Bayes-faktorimalleista paras on malli, jossa on viisi faktoria ja vakiotermi: hedelmällisyysluku, 20-24 -vuotiaiden, 25-29 -vuotiaiden, 30-34 -vuotiaiden ja 35-39 -vuotiaiden naisten lukumäärät. Muutokset näissä viidessä faktorissa riittävät selittämään 99,9 % muutoksista syntyvyyydessä (tarkemmin, varianssista).

Regressiokertoimien suuruudesta näkyy, että tärkeimmät faktorit ovat hedelmällisyysluku ja 25-29 sekä 30-34 -vuotiaiden naisten lukumäärä. Tämä vastaa tietoa, että synnyttäneiden keski-ikä on juuri tällä välillä (29 vuotta 1989 ja 30,7 vuotta 2016). Suurin osa synnyttäneistä on juuri 25-34 -vuotiaita.

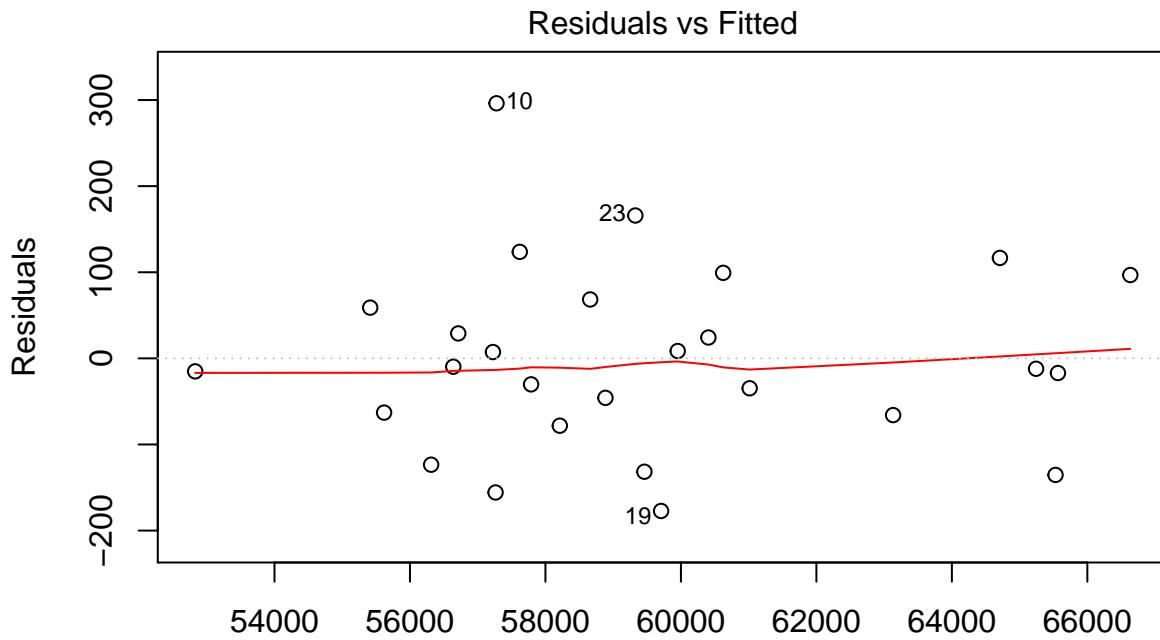
```
##
## Marginal Posterior Summaries of Coefficients:
##
## Using BMA
##
## Based on the top 256 models
##          post mean   post SD   post p(B != 0)
## Intercept      5.955e+04  2.308e+01  1.000e+00
## hedelmällisyysluku  3.391e+04  5.059e+02  1.000e+00
## naisten_lkm_15_19  9.915e-06  3.030e-03  1.847e-02
## naisten_lkm_20_24  3.933e-02  7.630e-03  9.865e-01
## naisten_lkm_25_29  1.116e-01  3.933e-03  1.000e+00
## naisten_lkm_30_34  1.053e-01  3.057e-03  1.000e+00
## naisten_lkm_35_39  6.181e-02  2.389e-03  1.000e+00
## naisten_lkm_40_45 -1.253e-04  1.076e-03  2.815e-02
## naisten_lkm_45_49 -1.281e-04  2.283e-03  3.348e-02
```

Erikoista kyllä, 40-44 ja 45-49-vuotiaiden naisten regressiokertoimet ovat negatiivisia, mikä vaikuttaa erikoiselta, mutta niiden faktorien todennäköisyys on vähäinen, juuri niinkuin intuitiokin sanoo.

Kun verrataan sovittetta aineistoon, näkyy että erot ovat datan ja mallin välillä erittäin pienet



Tarkastetaan vielä, millaisia residuaaleja jää selittämättä mallilla. Residuaalit käyttäytyvät hyvin, joten faktorimalli kuvailee syntyvyyden muutoksia.



Predictions under BMA

```
bas.lm(syntyneita ~ hedelmälisyytsluku + naisten_lkm_15_19 + naisten_lkm_20 ...
```

Havainnollisuuden vuoksi voi sanoa, että syntyneiden määärälle regressio ehdottaa yksinkertaisen mallin

$$S = 59550 + 33910H + 0.038N_{20-24} + 0.116N_{25-29} + 0.105N_{30-34} + 0.0618N_{35-39}$$

missä H on hedelmällisyytsluvun poikkeama vuosien 1989-2017 keskiarvosta ja N_{x-y}:t ovat x-y -ikäisten naisten lukumäärien poikkeamia vuosien 1989-2017 keskiarvosta. Tulkinta vakiotermille tässä yhteydessä on se, että se huomioi faktoreiden keskiarvojen vaikutuksen syntyneiden lukumäärään. Tämä regressiomalli ei tietenkäään huomioi epävarmuutta, joka näkyy täydellisessä Bayes-mallissa.

Pidempi aineisto

Yllä tarkasteltu malli toimi erinomaisesti vuosien 1989-2016 aineistolla. Siinä kuitenkin oletetaan että eri 5-vuotiskäryhmien painot säilyvät samoina koko aikavälin, vaikka tiedetään että synnyttäneiden keski-ikä on noussut.

Jotta nähdään, onko tällä vaikutusta, ajetaan sama malli pidemmälle aineistolle, vuosien 1970-2017 aineistolla.

```
syntyvyys_long = bas.lm(syntyneita ~ hedelmälisyytsluku + naisten_lkm_15_19 + naisten_lkm_20_24 + naisten_lkm_25_29, data = fertility_data, prior = "ZS-null", modelprior = uniform())
```

Nyt paras malli käyttää samoja faktoreita, mutta selittää hieman pienemmän osuuden varianssista, mutta kuitenkin 98,5 %. Ero ylläolevaan on se, että ajanjakson aikana synnyttäneiden keski-ikä on kasvanut, jolloin regressiokerrointenkin pitäisi muuttua vastaavasti. Tämä näkyy siinä, että tässä aineistossa regressiokerrointen painopiste on siirtynyt nuorempien ikiin, kuten odottaisikin.

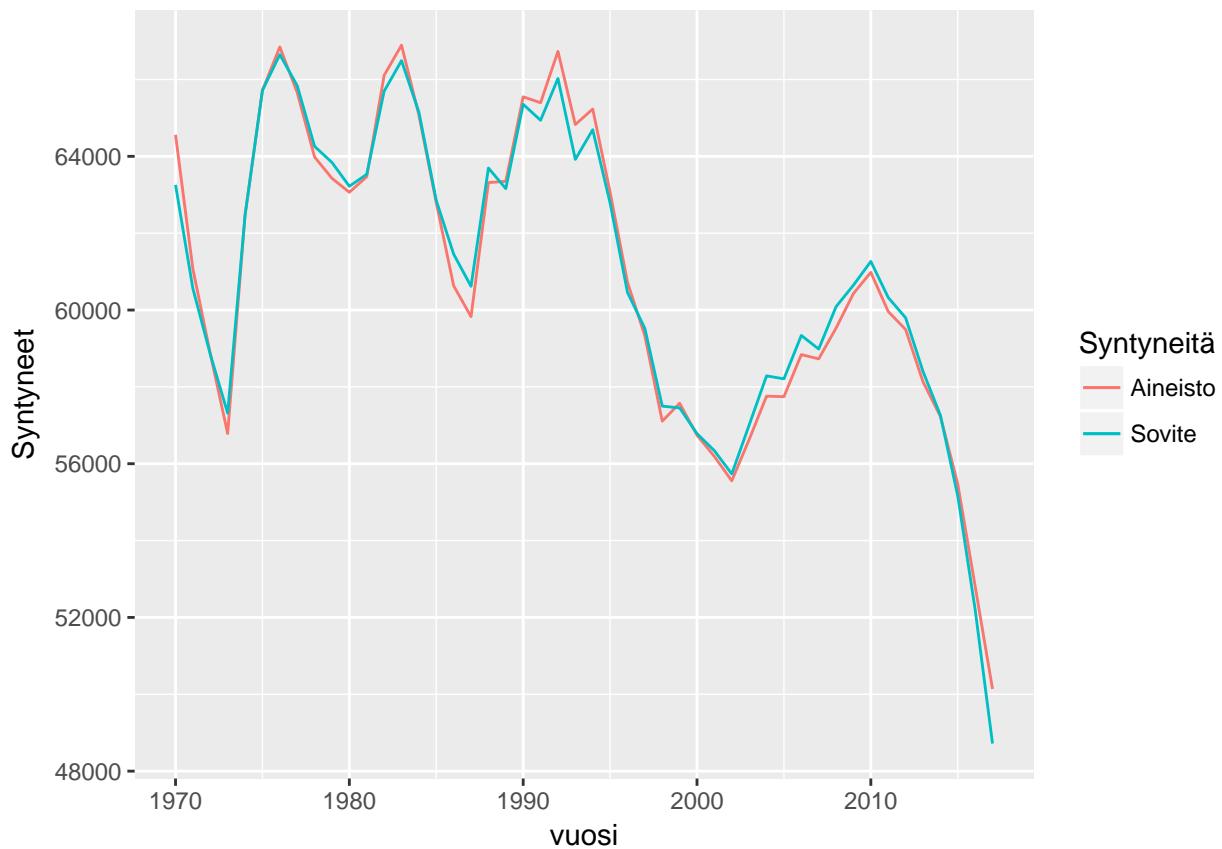
	P(B != 0 Y)	model 1	model 2	model 3	model 4	model 5
## Intercept		1.000	1.000	1.000	1.000	1.000
## hedelmälisyytsluku		1.000	1.000	1.000	1.000	1.000
## naisten_lkm_15_19		0.063	0.000	1.000	0.000	1.000
## naisten_lkm_20_24		1.000	1.000	1.000	1.000	1.000
## naisten_lkm_25_29		1.000	1.000	1.000	1.000	1.000

## naisten_lkm_30_34	1.000	1.000	1.000	1.000	1.000	1.000
## naisten_lkm_35_39	1.000	1.000	1.000	1.000	1.000	1.000
## naisten_lkm_40_45	0.948	1.000	1.000	0.000	0.000	1.000
## naisten_lkm_45_49	0.999	1.000	1.000	1.000	1.000	0.000
## BF	NA	1.000	0.066	0.055	0.003	0.001
## PostProbs	NA	0.888	0.059	0.049	0.003	0.001
## R2	NA	0.985	0.985	0.980	0.980	0.979
## dim	NA	8.000	9.000	7.000	8.000	8.000
## logmarg	NA	72.666	69.950	69.764	66.955	66.057

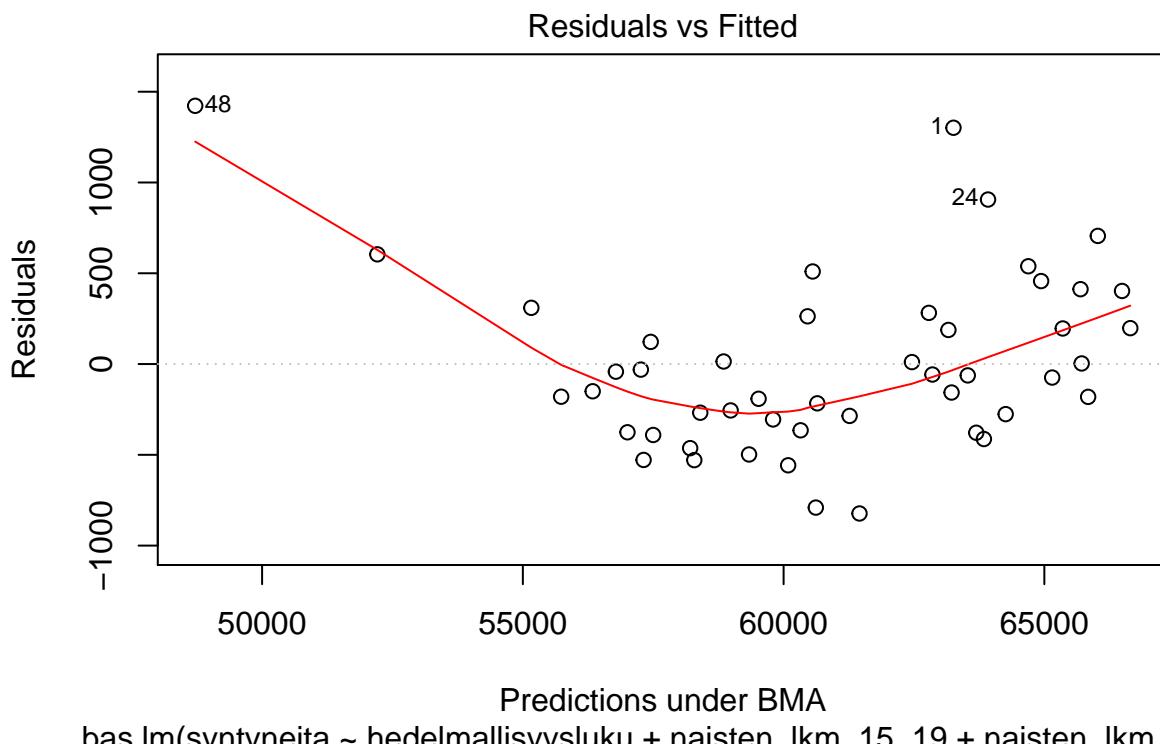
Regressiokertoimien suuruudesta näkyy, että nyt 20-24 ja 25-29 -vuotiaiden naisten lukumäärät vaikuttavat vahvemmin syntyneiden määrään. Tämä on järkevää, koska tarkasteluperiodin aikana synnyttäneiden keski-ikä on kasvanut.

```
##
## Marginal Posterior Summaries of Coefficients:
##
## Using BMA
##
## Based on the top 256 models
##          post mean  post SD  post p(B != 0)
## Intercept      6.087e+04 7.578e+01 1.000e+00
## hedelmallisyysluku 3.260e+04 1.080e+03 1.000e+00
## naisten_lkm_15_19 5.166e-04 4.136e-03 6.291e-02
## naisten_lkm_20_24 1.594e-01 1.167e-02 1.000e+00
## naisten_lkm_25_29 1.235e-01 6.938e-03 1.000e+00
## naisten_lkm_30_34 9.786e-02 6.968e-03 1.000e+00
## naisten_lkm_35_39 3.230e-02 6.637e-03 9.996e-01
## naisten_lkm_40_45 2.106e-02 7.667e-03 9.483e-01
## naisten_lkm_45_49 4.525e-02 8.385e-03 9.987e-01
```

Sovite on edelleenkin hyvä, mutta selviä eroja näkyy nyt faktorimallin ja data välillä.



Tarkastetaan vielä, millaisia residuaaleja jää selittämättä mallilla. Nyt residuaalit tasoa kuvaava punainen käyrä ei olekaan suora, mikä viittaa siihen että vuosien varrella tapahtuu muutos, jota tämä malli ei pysty selittämään.



Jälleen havainnollisuuden vuoksi syntyneiden määärälle regressio ehdottaa yksinkertaista mallia

$$S = 60870 + 32600H + 0.1594N_{20-24} + 0.1235N_{25-29} + 0.0987N_{30-34} + 0.03230N_{35-39} + 0.02106N_{40-44} + 0.04245N_{45-49}$$

missä H on hedelmällisyyslувun poikkeama vuosien 1970-2017 keskiarvosta ja $N_{\{x-y\}}$:t ovat x-y -ikäisten naisten lukumäärien poikkeamia vuosien 1970-2017 keskiarvosta. Vakiotermille huomioi faktoreiden keskiarvojen vaikutuksen syntyneiden lukumäärään.

Hedelmällisyyslувun malli selittävä näytävä tekijänä

Synnytysikäisten naisten lukumäärästä pystytään ennustamaan melko hyvin, eikä siihen vaikuta juurikaan muut tekijät, kuten vaikkapa työllisyys tai taloudellinen tilanne. Jos sellaiset tekijät vaikuttavat synnytyyteen, vaikutuksen pitäisi näkyä hedelmällisyyslувussa.

Yritetään sitten selittää hedelmällisyyslувun muutoksia käyttäen faktoreina kokoelmaa erilaisia tekijöitä ikäluokka- ja sukupuolikohtaisia avioituvuuksia ja pienituloisuusosuusasteita (60% mediaanitulosta); synnyttäjien keski-ikää, alle 1-vuotiaiden kuolleisuutta, eläkemenon bkt-osuutta, sekä sisäkaupungissa, ulkokaupungissa ja maaseudulla asuvien osuutta väestöstä.

```
syntyyvyys_reg = bas.lm(hedelmällisyysluku ~ pienituloisuus_18_24 + pienituloisuus_25_34 + lapsikuolleisuus
prior = "ZS-null", modelprior = uniform())
```

```
## Warning in bas.lm(hedelmällisyysluku ~ pienituloisuus_18_24 +
## pienituloisuus_25_34 + : dropping 1 rows due to missing data
```

Tuloksissa parhaana näkyy kolmen faktorin ja vakiotermin malli, jossa faktoreina ovat 25-29-vuotiaiden avioituvuuus, sisäkaupungissa ja muualla kaupungissa asuvien osuus. Näillä faktoreilla on mahdollista selittää noin 92 % hedelmällisyyslувun varianssista datassa. Hieman yllättäen synnyttäjien keski-ikä ei näy merkittävä näytävä tekijänä.

	P(B != 0 Y)	model 1	model 2	model 3
## Intercept	1.000	1.000	1.000	1.000
## pienituloisuus_18_24	0.206	0.000	0.000	0.000
## pienituloisuus_25_34	0.170	0.000	0.000	0.000
## lapsikuolleisuus	0.146	0.000	0.000	0.000
## synnyttäjien_keskiika	0.287	0.000	0.000	0.000
## avioituvuuus_naiset_25_29_lag1	0.591	1.000	0.000	0.000
## avioituvuuus_naiset_30_34_lag1	0.537	0.000	1.000	1.000
## elakemeno	0.218	0.000	0.000	0.000
## kaupungissaasuvat	0.679	1.000	0.000	1.000
## maaseudulla	0.232	0.000	0.000	0.000
## sisäkaupungissa	0.923	1.000	1.000	1.000
## BF	NA	1.000	0.373	0.208
## PostProbs	NA	0.160	0.060	0.033
## R2	NA	0.922	0.896	0.910
## dim	NA	4.000	3.000	4.000
## logmarg	NA	23.368	22.383	21.799
##		model 4	model 5	
## Intercept	1.000	1.000		
## pienituloisuus_18_24	0.000	0.000		
## pienituloisuus_25_34	0.000	0.000		
## lapsikuolleisuus	0.000	0.000		
## synnyttäjien_keskiika	0.000	0.000		
## avioituvuuus_naiset_25_29_lag1	1.000	1.000		
## avioituvuuus_naiset_30_34_lag1	0.000	0.000		

```

## elakemeno           1.000  0.000
## kaupungissaasuvat 1.000  1.000
## maaseudulla        0.000  1.000
## sisakaupungissa    1.000  1.000
## BF                 0.201  0.180
## PostProbs          0.032  0.029
## R2                 0.925  0.924
## dim                5.000  5.000
## logmarg            21.764 21.652

```

Todennäköisimmässä mallissa regressiokertoimet ovat

```

##
## Marginal Posterior Summaries of Coefficients:
##
## Using HPM
##
## Based on the top 1 models
##                                         post mean  post SD  post p(B != 0)
## Intercept                         1.776296  0.003982 1.000000
## pienituloisuus_18_24             0.000000  0.000000 0.205574
## pienituloisuus_25_34             0.000000  0.000000 0.169991
## lapsikuolleisuus                0.000000  0.000000 0.145603
## synnyttajien_keskiika          0.000000  0.000000 0.287413
## avioituvuus_naiset_25_29_lag1  0.007458  0.001022 0.590711
## avioituvuus_naiset_30_34_lag1  0.000000  0.000000 0.536778
## elakemeno                       0.000000  0.000000 0.218424
## kaupungissaasuvat              0.074730  0.008616 0.678634
## maaseudulla                     0.000000  0.000000 0.232215
## sisakaupungissa                -0.151944 0.025115 0.923196

```

Regressiokertoimista näkyy, että avioituvuus lisää syntyyvyyttä, mutta matallisesti: yhden prosenttiyksikön nousu avioituvuudessa kytkeytyisi hedelmällisyyslувun kasvuun 1,49:stä 1,494:een. Viivästetty avioituvuus on laskenut 15,9 prosenttiyksikköä 25-29-vuotiaissa vuodesta 2010 vuoteen 2016, jolloin mallin mukaan sillä voi selittää 0,119:n verran hedelmällisyyslувun pudotuksesta.

Kaupungistuminen liittyy jälleen regression mukaan hedelmällisyyslukuun: sisäkaupungissa asuvien osuuden kasvu alentaa hedelmällisyyslukua (kerroin -0.151944) ja muualla kaupungissa asuvien osuuden kasvu kasvattaa hedelmällisyyslukua (kerroin 0.074730). Tiedetään, että Suomen suurissa kaupungeissa on alimmat hedelmällisyyslувut, joten tulos on uskottava. Vuodesta 2010 vuoteen 2016 muualla kaupungissa asuvien osuus on kasvanut 1,8 prosenttiyksikkö ja sisäkaupungissa asuvien osuus on kasvanut 1,2 prosenttiyksikkö. Näiden yhteisvaikutus hedelmällisyyslukuun on lasku 0,048:n verran.

Mallin mukaan vuosien 2010 ja 2016 välisestä hedelmällisyyslувun 0,30:n pudotuksesta nämä kolme faktorit selittävät hieman yli puolet (0,166).

Aggregaattikertoimet kaikissa mallissa todennäköisyyskäytäntöillä painotettuna ovat

```

##
## Marginal Posterior Summaries of Coefficients:
##
## Using BMA
##
## Based on the top 3 models
##                                         post mean  post SD  post p(B != 0)
## Intercept                         1.776296  0.004145 1.000000
## pienituloisuus_18_24             0.000000  0.000000 0.205574

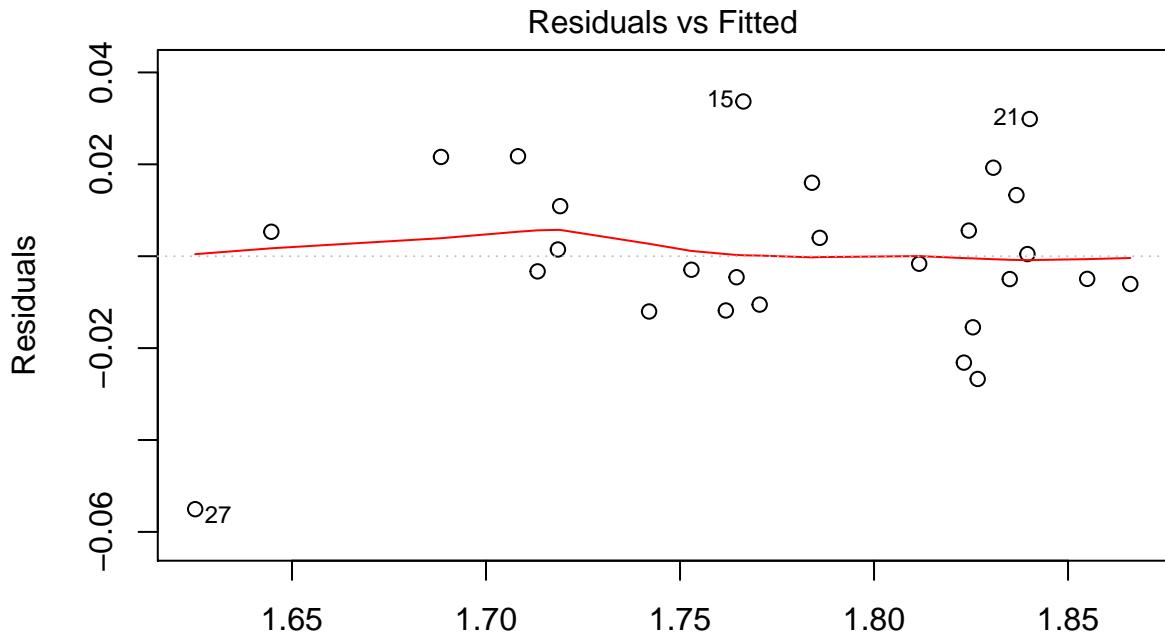
```

```

## pienituloisuus_25_34          0.000000  0.000000  0.169991
## lapsikuolleisuus            0.000000  0.000000  0.145603
## synnyttajien_keskiika       0.000000  0.000000  0.287413
## avioituvuus_naiset_25_29_lag1 0.004715  0.003687  0.590711
## avioituvuus_naiset_30_34_lag1 0.002739  0.003647  0.536778
## elakemeno                   0.000000  0.000000  0.218424
## kaupungissaasuvat           0.050609  0.033583  0.678634
## maaseudulla                 0.000000  0.000000  0.232215
## sisakaupungissa              -0.130177  0.039746  0.923196

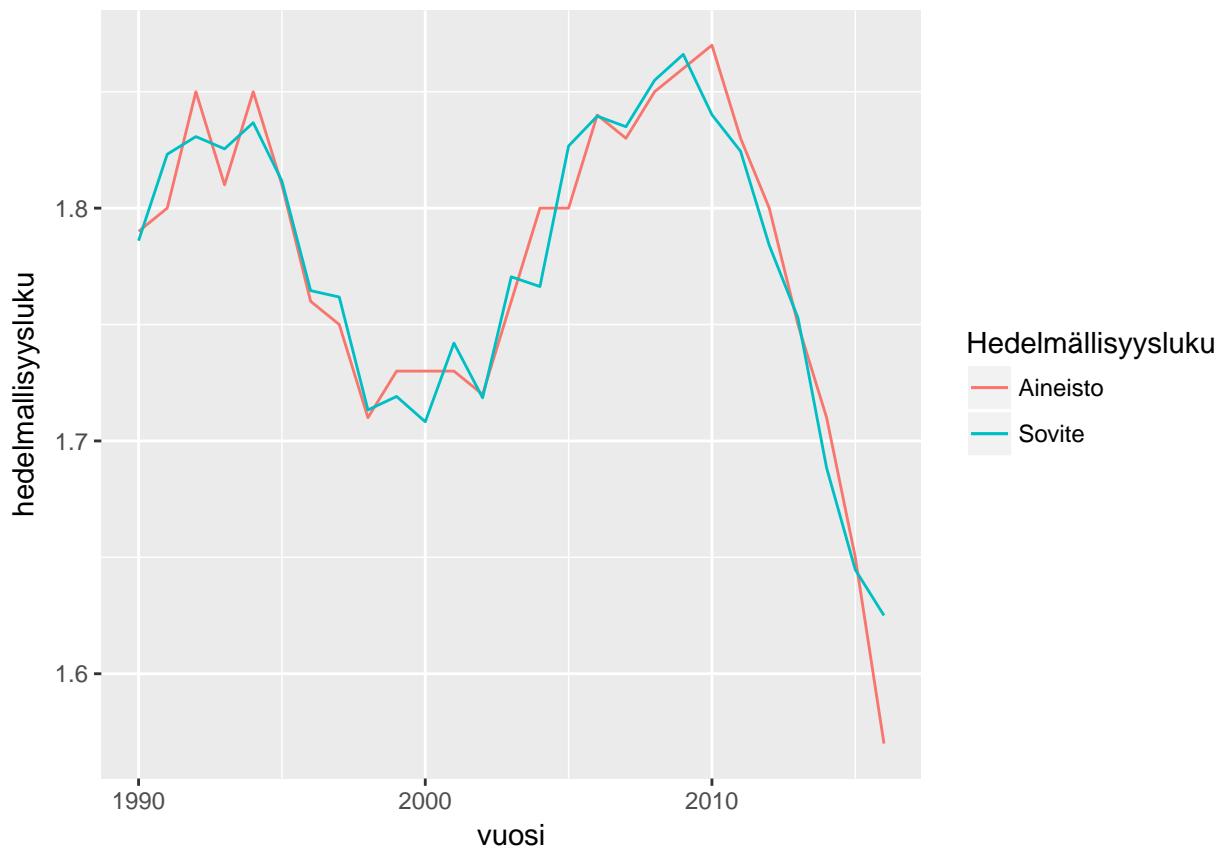
```

Tarkastetaan vielä, millaisia residuaaleja jää selittämättä mallilla



Predictions under BMA
bas.lm(hedelmallisysluku ~ pienituloisuus_18_24 + pienituloisuus_25_34 + I ...

Ja lopuksi verrataan mallin sovitetta



Residuaaleista näkee, että malli selittää heikoinmin viimeisimpia havaintoja. Aiemmin laskettiin, että juuri tällä aikavälillä faktorit selittivät vain hieman yli puolet varianssista.⁸

Työttömyys ja pienituloisuus ja hedelmällisyysluvun muutokset

Yritetään sitten selittää hedelmällisyysluvun muutoksia käyttäen faktoreina 10-vuotisikäluokka- ja sukupuolikohtaisia työttömien osuutta kohortista, työttömyysasteita ja työvoimanulkopuolisuuksasteita(viivästettynä) ja pienituloisuusosuusasteita (60% mediaanitulosta), synnyttäjien keski-ikää, avioituvuutta, eläkemenon osuutta BKT:stä sekä ulommassa kaupungissa, maaseudulla ja sisemmässä kaupungissa asuvien osuksia.

```
syntyvyys_reg = bas.lm(hedelmällisyysluku ~ pienituloisuus_18_24 + pienituloisuus_25_34 +
lapsikuolleisuus + työttömyys_per_vaesto_miehet_25_34_lag1 +työttömyys_per_vaesto_miehet_35_44_lag1 +
+ työttömyys_per_vaesto_naiset_25_34_lag1 +työttömyys_per_vaesto_naiset_35_44_lag1 +
tyollisyysaste_miehet_25_34_lag1 +tyollisyysaste_naiset_25_34_lag1
+ tyollisyysaste_miehet_35_44_lag1 +tyollisyysaste_naiset_35_44_lag1
+ulkopuolistet_naiset_25_34_lag1 +ulkopuolistet_miehet_25_34_lag1 +synnyttajien_keskiika + avioituvuus_n
prior = "ZS-null", modelprior = uniform())
```

```
## Warning in bas.lm(hedelmällisyysluku ~ pienituloisuus_18_24 +
## pienituloisuus_25_34 + : dropping 1 rows due to missing data
```

Tuloksista näkyy selvästi, että sisemmässä kaupungissa ja maaseudulla asuvien osuus väestöstä, synnyttäjien keski-ikä liittyy kytkeytystä selvästi hedelmällisyyslukoon. Muita vaikuttavia, mutta vähemmän merkittäviä tekijöitä ovat synnyttäjien keski-ikä,

Ja kun katsomme tuloksia, näkyy malli, jossa faktoreina ovat 25-29 -vuotiaiden naisten avioituvuus, sisemmässä kaupungissa asuvien osuus ja ulommassa kaupungissa asuvien osuus. Nämä kolmella faktoreilla on mahdollista selittää noin 92 % syntyneiden lukumäärän varianssista datassa.

	P(B != 0 Y)	model 1	model 2	
## Intercept		1.000	1.000	
## pienituloisuus_18_24	0.406	0.000	0.000	
## pienituloisuus_25_34	0.166	0.000	0.000	
## lapsikuolleisuus	0.153	0.000	0.000	
## tyottomyys_per_vaesto_miehet_25_34_lag1	0.361	0.000	1.000	
## tyottomyys_per_vaesto_miehet_35_44_lag1	0.286	0.000	0.000	
## tyottomyys_per_vaesto_naiset_25_34_lag1	0.234	0.000	0.000	
## tyottomyys_per_vaesto_naiset_35_44_lag1	0.203	0.000	0.000	
## tyollisyysaste_miehet_25_34_lag1	0.260	0.000	0.000	
## tyollisyysaste_naiset_25_34_lag1	0.226	0.000	0.000	
## tyollisyysaste_miehet_35_44_lag1	0.232	0.000	0.000	
## tyollisyysaste_naiset_35_44_lag1	0.230	0.000	0.000	
## ulkopuolistet_naiset_25_34_lag1	0.155	0.000	0.000	
## ulkopuolistet_miehet_25_34_lag1	0.175	0.000	0.000	
## synnyttajien_keskiika	0.675	0.000	1.000	
## avioituvuus_naiset_25_29_lag1	0.295	1.000	0.000	
## avioituvuus_naiset_30_34_lag1	0.851	0.000	1.000	
## elakemeno	0.216	0.000	0.000	
## kaupungissaasuvat	0.402	1.000	1.000	
## maaseudulla	0.272	0.000	0.000	
## sisakaupungissa	0.408	1.000	0.000	
## BF		NA	1.000	
## PostProbs		NA	0.007	
## R2		NA	0.922	
## dim		NA	4.000	
## logmarg		NA	23.368	
##			22.719	
		model 3	model 4	model 5
## Intercept		1.000	1.000	1.000
## pienituloisuus_18_24	0.000	0.000	0.000	
## pienituloisuus_25_34	0.000	0.000	0.000	
## lapsikuolleisuus	0.000	0.000	0.000	
## tyottomyys_per_vaesto_miehet_25_34_lag1	0.000	0.000	0.000	
## tyottomyys_per_vaesto_miehet_35_44_lag1	1.000	0.000	0.000	
## tyottomyys_per_vaesto_naiset_25_34_lag1	0.000	0.000	0.000	
## tyottomyys_per_vaesto_naiset_35_44_lag1	0.000	0.000	0.000	
## tyollisyysaste_miehet_25_34_lag1	0.000	1.000	1.000	
## tyollisyysaste_naiset_25_34_lag1	0.000	0.000	0.000	
## tyollisyysaste_miehet_35_44_lag1	0.000	0.000	0.000	
## tyollisyysaste_naiset_35_44_lag1	0.000	0.000	0.000	
## ulkopuolistet_naiset_25_34_lag1	0.000	0.000	0.000	
## ulkopuolistet_miehet_25_34_lag1	0.000	0.000	0.000	
## synnyttajien_keskiika	1.000	1.000	1.000	
## avioituvuus_naiset_25_29_lag1	0.000	0.000	0.000	
## avioituvuus_naiset_30_34_lag1	1.000	1.000	1.000	
## elakemeno	0.000	0.000	0.000	
## kaupungissaasuvat	1.000	0.000	1.000	
## maaseudulla	0.000	0.000	0.000	
## sisakaupungissa	0.000	0.000	0.000	
## BF	0.430	0.418	0.389	
## PostProbs	0.003	0.003	0.003	
## R2	0.930	0.915	0.930	
## dim	5.000	4.000	5.000	
## logmarg	22.525	22.496	22.425	

Viivästetyt naisten työttömyysasteet 25-34 ja 35-44 vuotiailla kytkeytyvät hedelmällisyyslukuun, mutta melko pienellä todennäköisyydellä. Bayesilainen keskiarvoregressiokertoimet ovat -0,0015 ja -0,00055. Naisten työssömyysasteen kasvaessa hedelmällisyysluku siis pienenee, mutta hyvin matallisesti: yhden prosenttiyksikön kasvu työttömyysasteessa alentaisi hedelmällisyyslukua vain 0,0015. Se, että vastaavien suureiden etumerkkien erilaisuus viittaa mahdolliseen multikollineaarisuusongelmaan. Regressiokerrointen mukaan 15-24 -vuotiaiden työllisyysasteen kasvu alentaisi hedelmällisyyslukua hienoistesti, kun taas 25-34 -vuotiaiden työllisyysasteen kasvu nostaisi hieman hedelmällisyyslukua. On mahdollista, että kertoimet toisiaan menevät näin, mutta eivät juurikaan kerro voimakkaasta yhteydestä työllisyysasteen ja syntynytyden välillä. Miesten työllisyysasteet eivät juurikaan kytkeydy työllisyteen.

Alle 18-vuotiaiden pienituloisuus liittyy hedelmällisyyslukuun, mutta regressiokerroin on vain -0,02. Tällöin 1 %-yksikön alenemaa alle 18-vuotiaiden pienituloisuudessa liittyy hedelmällisyysluvun kasvuun 1,49:stä 1,51:een. Myös 50-64 -vuotiaiden pienituloisuus liittyy hedelmällisyyslukuun, mutta regressiokerroin on +0,0089. Positiivinen etumerkki viittaa tässä tapauksessa multikollineaarisuusongelmaan, mutta jos uskomme kertoimen olevan oikein, liittyy 1 %-yksikön pieneminen 50-64 -vuotiaiden pienituloisuudessa hedelmällisyysluvun alenemiseen 1,49:stä 1,48:aan. Kyse ei siis olisi isoäiti-efektistä, jossa isovanhempien apu lastenhoidossa lisäisi syntynyttää. Tämä kerroin ei vaikuta uskottavalta.

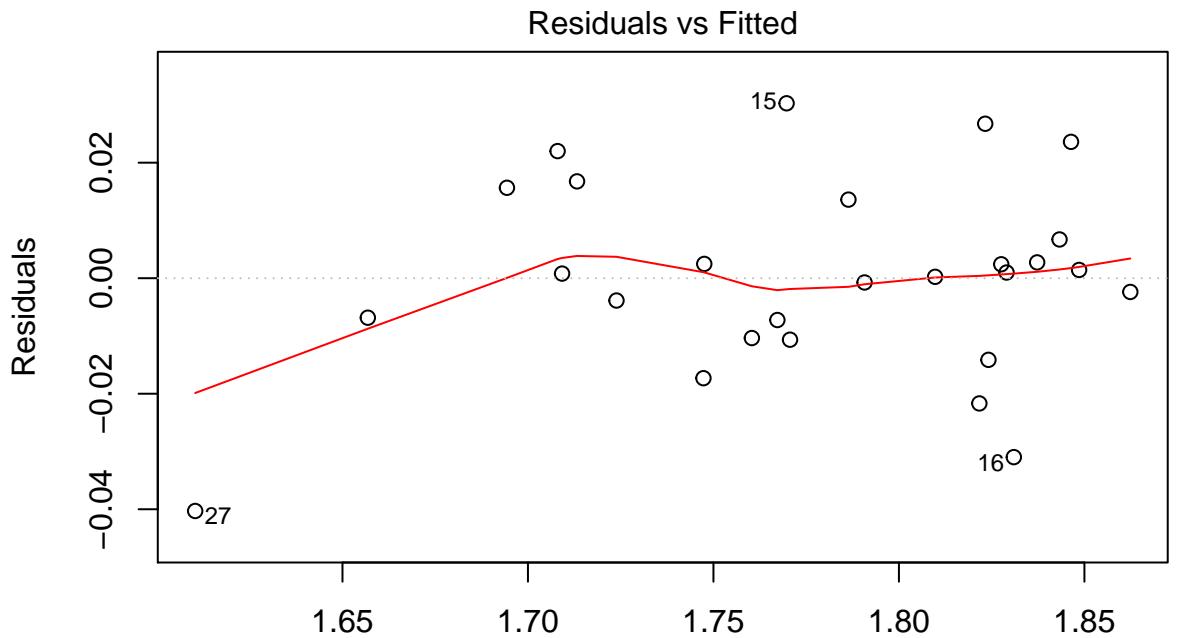
Sisemmässä kaupungissa asuvien osuus liittyy regression mukaan hedelmällisyyslukuun kertoimella -0,039. Tällöin 1 %-yksikön kasvu sisemmässä kaupungissa asuvien osuudessa alentaisi hedelmällisyyslukua 1,49:stä 1,44:een. Tiedetään, että Suomen suurissa kaupungeissa on alimmat hedelmällisyysluvut, joten tulos on uskottava.

Synnyttäjien keski-ikä kytkeytyy vahvasti hedelmällisyyslukuun (-0,15). Jos synnyttäjien keski-ikä alenisi yhden vuoden, kasvaisi hedelmällisyysluku 1,49:stä 1,64:ään, jos arvio regressiokertoimesta on oikea.

```
plot(syntynyys_reg, subset=c(4:13), ask=F)

## Warning in plot.window(...): "subset" is not a graphical parameter
## Warning in plot.xy(xy, type, ...): "subset" is not a graphical parameter
## Warning in axis(side = side, at = at, labels = labels, ...): "subset" is
## not a graphical parameter

## Warning in axis(side = side, at = at, labels = labels, ...): "subset" is
## not a graphical parameter
## Warning in box(...): "subset" is not a graphical parameter
## Warning in title(...): "subset" is not a graphical parameter
## Warning in plot.xy(xy.coords(x, y), type = type, ...): "subset" is not a
## graphical parameter
## Warning in title(sub = sub.caption, ...): "subset" is not a graphical
## parameter
```

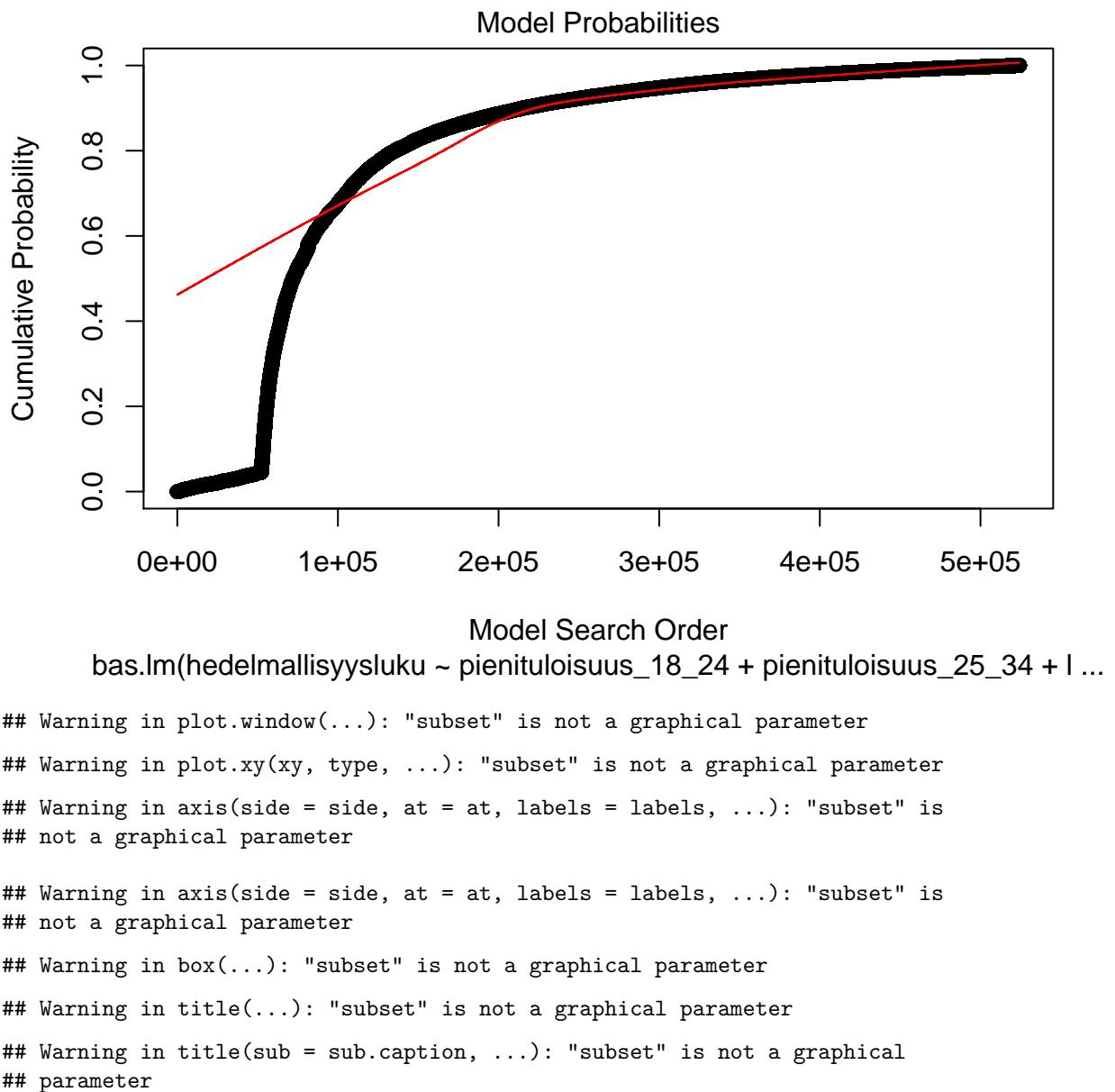


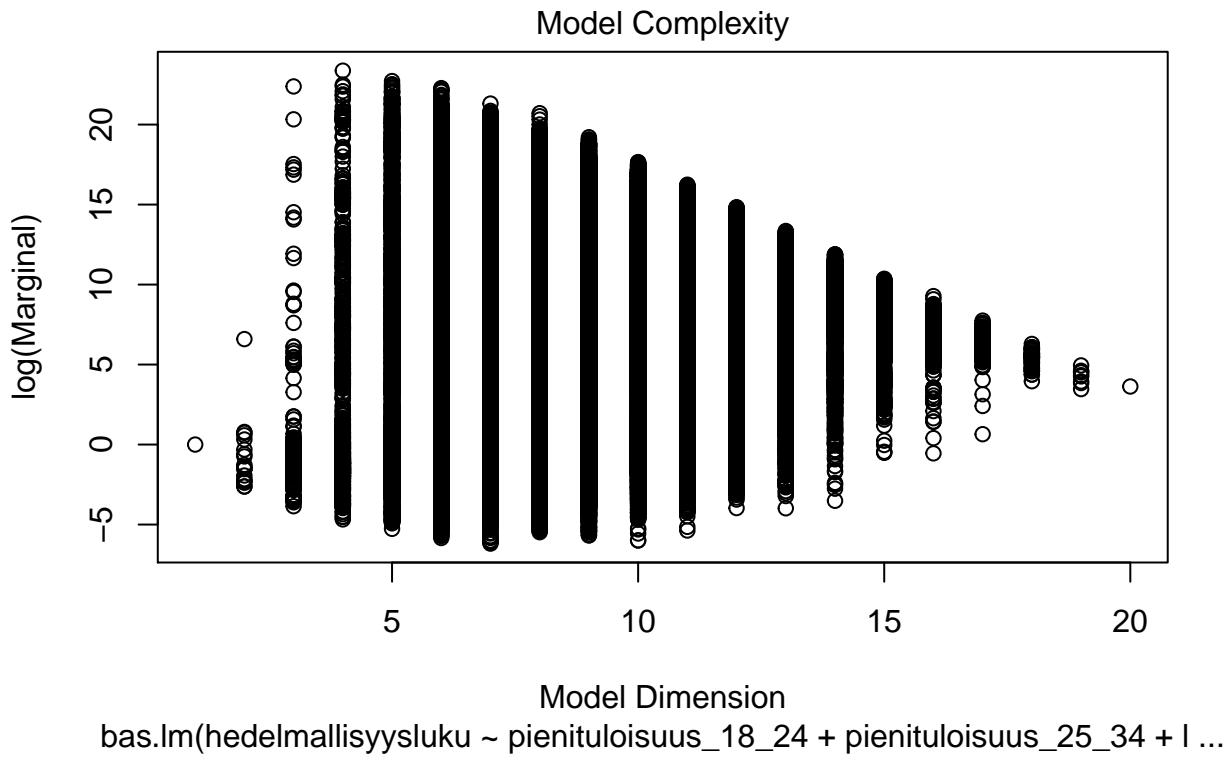
Predictions under BMA

```
bas.lm(hedelmallisyysluku ~ pienituloisuus_18_24 + pienituloisuus_25_34 + I ...
```

```
## Warning in plot.window(...): "subset" is not a graphical parameter
## Warning in plot.xy(xy, type, ...): "subset" is not a graphical parameter
## Warning in axis(side = side, at = at, labels = labels, ...): "subset" is
## not a graphical parameter

## Warning in axis(side = side, at = at, labels = labels, ...): "subset" is
## not a graphical parameter
## Warning in box(...): "subset" is not a graphical parameter
## Warning in title(...): "subset" is not a graphical parameter
## Warning in title(sub = sub.caption, ...): "subset" is not a graphical
## parameter
```

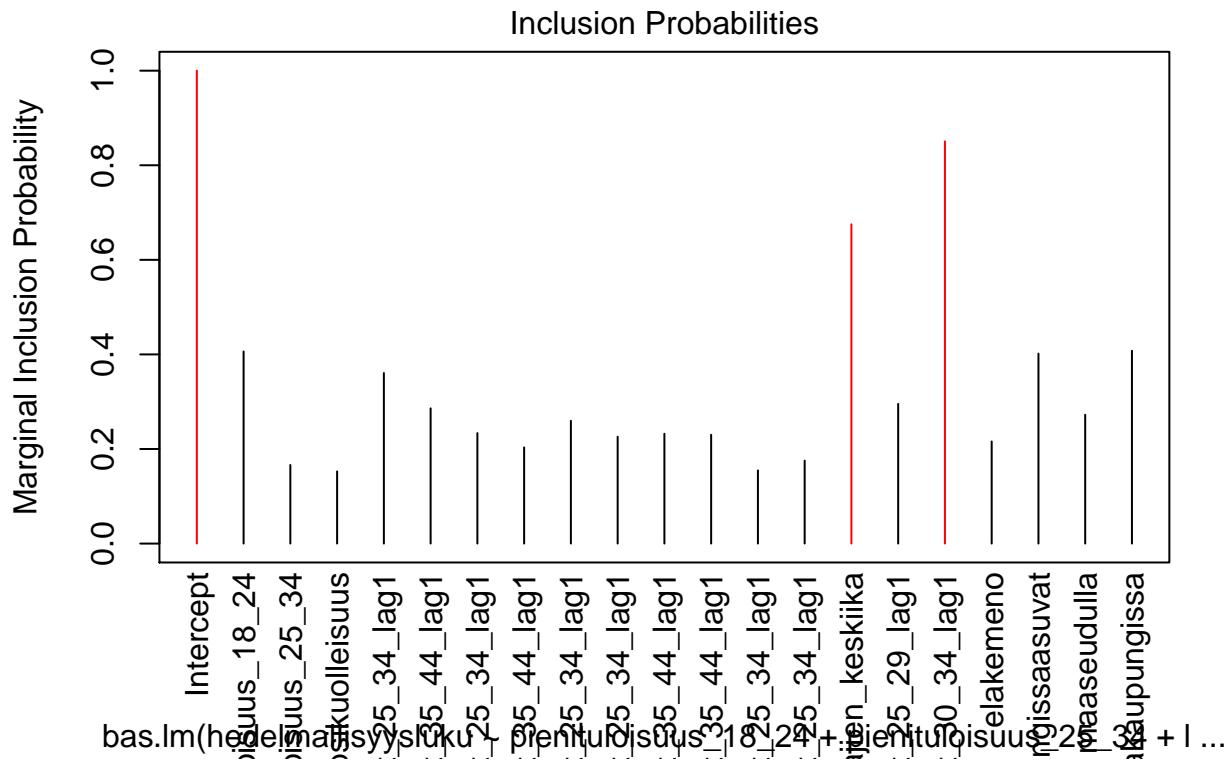




```
bas.lm(hedelmallisyysluku ~ pienituloisuus_18_24 + pienituloisuus_25_34 + I ...
```

```
## Warning in plot.window(...): "subset" is not a graphical parameter
## Warning in plot.xy(xy, type, ...): "subset" is not a graphical parameter
## Warning in axis(side = side, at = at, labels = labels, ...): "subset" is
## not a graphical parameter

## Warning in axis(side = side, at = at, labels = labels, ...): "subset" is
## not a graphical parameter
## Warning in box(...): "subset" is not a graphical parameter
## Warning in title(...): "subset" is not a graphical parameter
## Warning in title(sub = sub.caption, ...): "subset" is not a graphical
## parameter
## Warning in mtext(x$namesx, side = 1, line = 0.25, at = variables, las =
## 2, : "subset" is not a graphical parameter
```

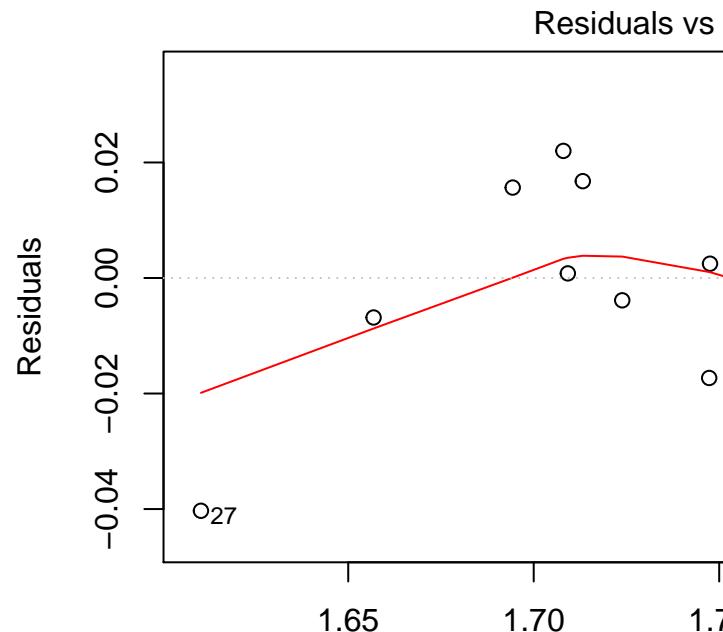


```
##  
## Marginal Posterior Summaries of Coefficients:  
##  
## Using BMA  
##  
## Based on the top 524288 models  
##  
## Intercept post mean post SD  
## 1.776e+00 4.070e-03  
## pienituloisuus_18_24 1.801e-03 2.752e-03  
## pienituloisuus_25_34 -1.567e-04 3.049e-03  
## lapsikuolleisuus 1.160e+00 8.670e+00  
## tyottomyys_per_vaesto_miehet_25_34_lag1 4.487e-03 1.069e-02  
## tyottomyys_per_vaesto_miehet_35_44_lag1 2.814e-03 9.283e-03  
## tyottomyys_per_vaesto_naiset_25_34_lag1 -1.488e-03 5.359e-03  
## tyottomyys_per_vaesto_naiset_35_44_lag1 -5.537e-04 4.893e-03  
## tyollisyysaste_miehet_25_34_lag1 1.344e-04 8.266e-03  
## tyollisyysaste_naiset_25_34_lag1 7.632e-04 4.234e-03  
## tyollisyysaste_miehet_35_44_lag1 9.168e-04 6.280e-03  
## tyollisyysaste_naiset_35_44_lag1 1.434e-03 4.414e-03  
## ulkopuolistet_naiset_25_34_lag1 -3.101e-05 3.764e-03  
## ulkopuolistet_miehet_25_34_lag1 1.227e-03 8.451e-03  
## synnyttajien_keskiika -1.460e-01 1.288e-01  
## avioituvuus_naiset_25_29_lag1 1.099e-03 2.847e-03  
## avioituvuus_naiset_30_34_lag1 5.997e-03 3.026e-03  
## elakemeno -1.269e-03 5.249e-03  
## kaupungissaasuvat 2.104e-02 3.772e-02  
## maaseudulla 1.646e-03 3.345e-02  
## sisakaupungissa -3.936e-02 6.563e-02  
##  
## Intercept post p(B != 0)  
## 1.000e+00
```

```

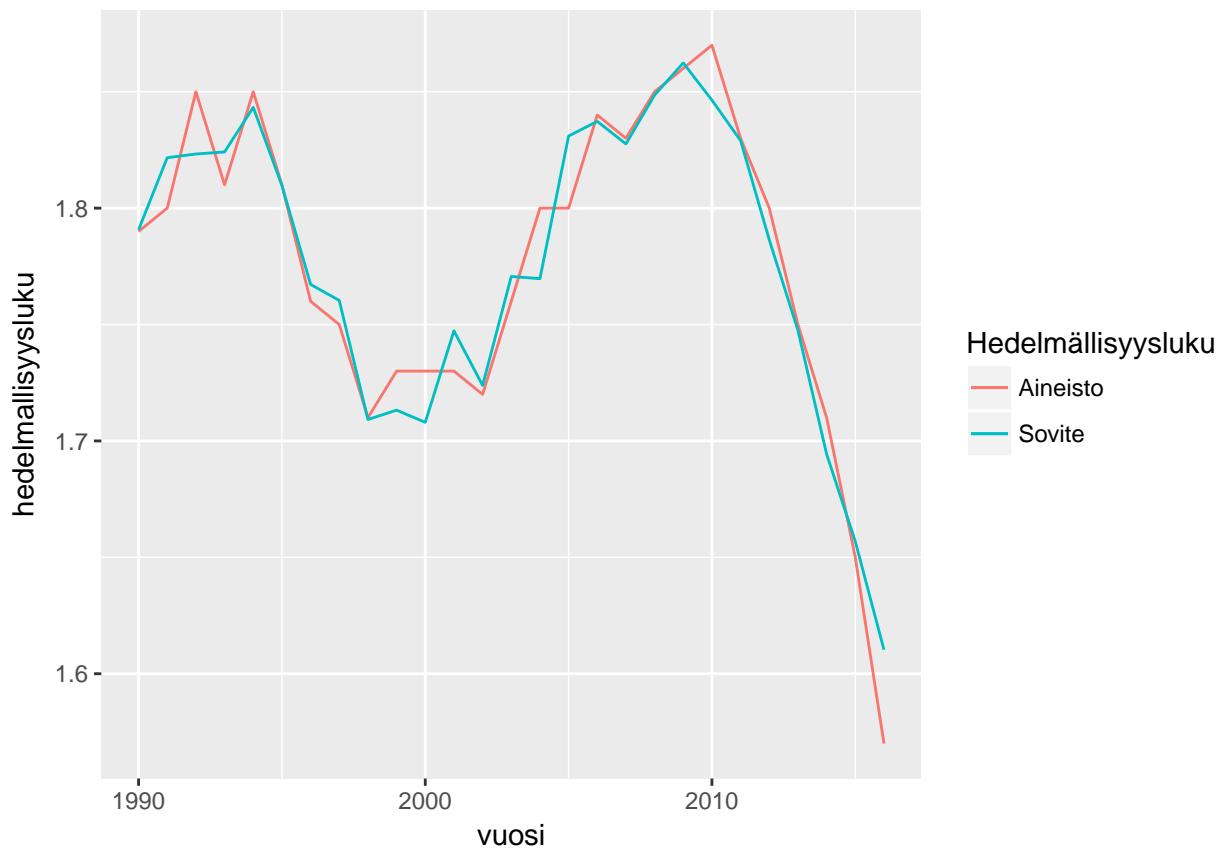
## pienituloisuus_18_24          4.062e-01
## pienituloisuus_25_34          1.662e-01
## lapsikuolleisuus             1.527e-01
## työttömyys_per_vaesto_miehet_25_34_lag1 3.607e-01
## työttömyys_per_vaesto_miehet_35_44_lag1 2.860e-01
## työttömyys_per_vaesto_naiset_25_34_lag1 2.337e-01
## työttömyys_per_vaesto_naiset_35_44_lag1 2.035e-01
## työllisyysaste_miehet_25_34_lag1      2.597e-01
## työllisyysaste_naiset_25_34_lag1       2.260e-01
## työllisyysaste_miehet_35_44_lag1       2.323e-01
## työllisyysaste_naiset_35_44_lag1       2.301e-01
## ulkopuoliset_naiset_25_34_lag1        1.547e-01
## ulkopuoliset_miehet_25_34_lag1         1.755e-01
## synnyttajien_keskiika                6.752e-01
## avioituvuus_naiset_25_29_lag1         2.954e-01
## avioituvuus_naiset_30_34_lag1          8.505e-01
## elakemeno                           2.160e-01
## kaupungissaasuvat                   4.017e-01
## maaseudulla                         2.725e-01
## sisakaupungissa                      4.079e-01

```



Tarkastetaan vielä, millaisia residuaaleja jää selittämättä mallilla
Ja lopuksi verrataan mallin sovitetta

Predictions und
bas.lm(hedelmallisyysluku ~ pienituloisuus_



Synnyttäjien keski-ikään vaikuttavat faktorit

Yritetään sitten selittää muutoksia synnyttäjen keski-iässä käyttäen faktoreina 24-39 -vuotiaiden naisten työllisyyssastetta, ikäluokka- ja sukupuolikohtaisia pienituloisuusosuusasteita (60% mediaanitulosta), kaupungistumista kuvaavia tekijöitä, eläkemeno osuutta BKT:stä, lapsikuolleisuutta.

Ja kun katsomme tuloksia, parhaana näkyy malli, jossa faktoreina ovat 25-34 -vuotiaiden naisten työllisyyssaste ja kaupungistumisesta kertovat tekijät. Näillä neljällä faktoreilla on mahdollista selittää 99,4 % syntyneiden lukumäärän varianssista datassa.

	P(B != 0 Y)	model 1	model 2	model 3	
## Intercept		1.000	1.000	1.000	
## pienituloisuus_25_34	0.321	0.000	0.000	1.000	
## lapsikuolleisuus	0.059	0.000	0.000	0.000	
## työllisyyssaste_naiset_25_34_lag1	1.000	1.000	1.000	1.000	
## elakemeno	0.125	0.000	0.000	0.000	
## kaupungissaasuvat	0.606	1.000	0.000	1.000	
## maaseudulla	0.908	1.000	1.000	1.000	
## sisakaupungissa	0.999	1.000	1.000	1.000	
## BF		NA	1.000	0.645	0.582
## PostProbs		NA	0.324	0.209	0.188
## R2		NA	0.994	0.991	0.995
## dim		NA	5.000	4.000	6.000
## logmarg		NA	47.532	47.093	46.991
## Intercept		1.000	1.000		

```

## pienituloisuus_25_34          1.000  0.000
## lapsikuolleisuus             0.000  0.000
## työllisyysaste_naiset_25_34_lag1 1.000  1.000
## elakemeno                      0.000  1.000
## kaupungissaasuvat              0.000  0.000
## maaseudulla                     1.000  1.000
## sisakaupungissa                 1.000  1.000
## BF                            0.200  0.085
## PostProbs                      0.065  0.028
## R2                            0.992  0.992
## dim                           5.000  5.000
## logmarg                         45.920 45.066

```

Regressiokertoimien suuruudesta näkyy, että 25-34 -vuotiaiden naisten työllisyysaste alentaa keski-ikää. Regressiokerroin on -0,02, jolloin 1 prosenttiyksikön kasvu työllisyysasteessa alentaisi keski-ikää 0,02 vuotta. Maaseudulla asuvien osuuden kasvu yhdellä prosenttiyksiköllä alentaisi keski-ikää 0,18 vuotta, kun taas sisemmässä kaupungissa asuvien osuuden kasvu yhdellä prosenttiyksiköllä kasvattaisi keski-ikää 0,26 vuotta.

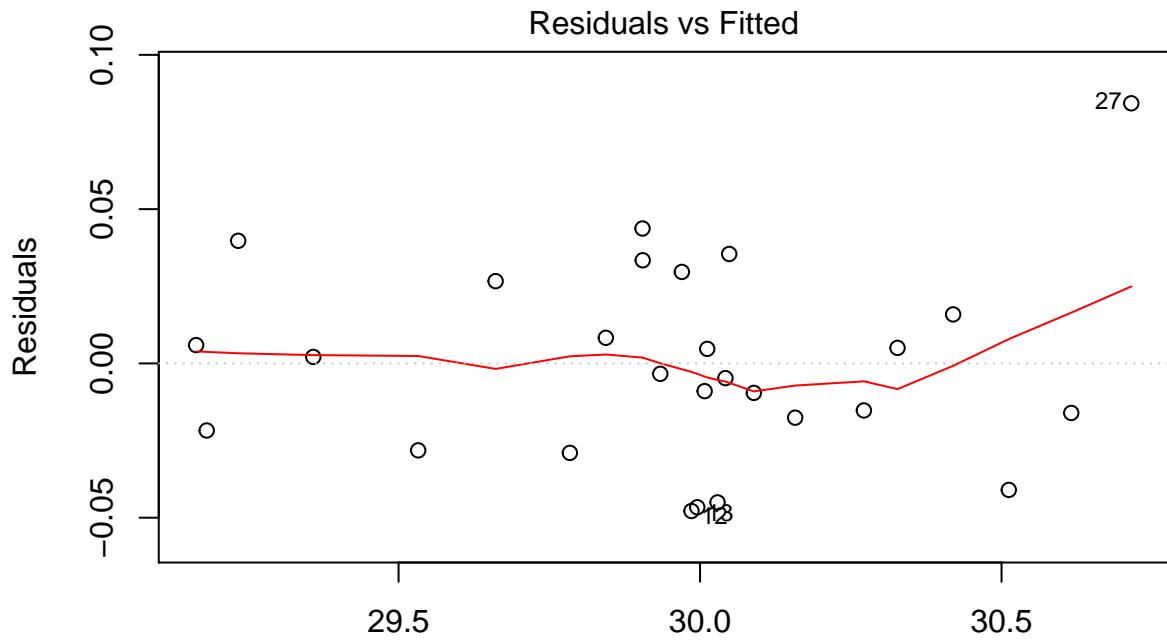
```

##
## Marginal Posterior Summaries of Coefficients:
##
## Using BMA
##
## Based on the top 128 models
##                                         post mean  post SD  post p(B != 0)
## Intercept                           29.951947  0.007263  1.000000
## pienituloisuus_25_34            -0.007342  0.012013  0.321091
## lapsikuolleisuus                  -0.033297 11.434603  0.059381
## työllisyysaste_naiset_25_34_lag1 -0.020996  0.002193  0.999998
## elakemeno                          0.001773  0.005979  0.124673
## kaupungissaasuvat                 -0.106313  0.110908  0.605654
## maaseudulla                        -0.177006  0.120224  0.907766
## sisakaupungissa                   0.258021  0.061956  0.999089

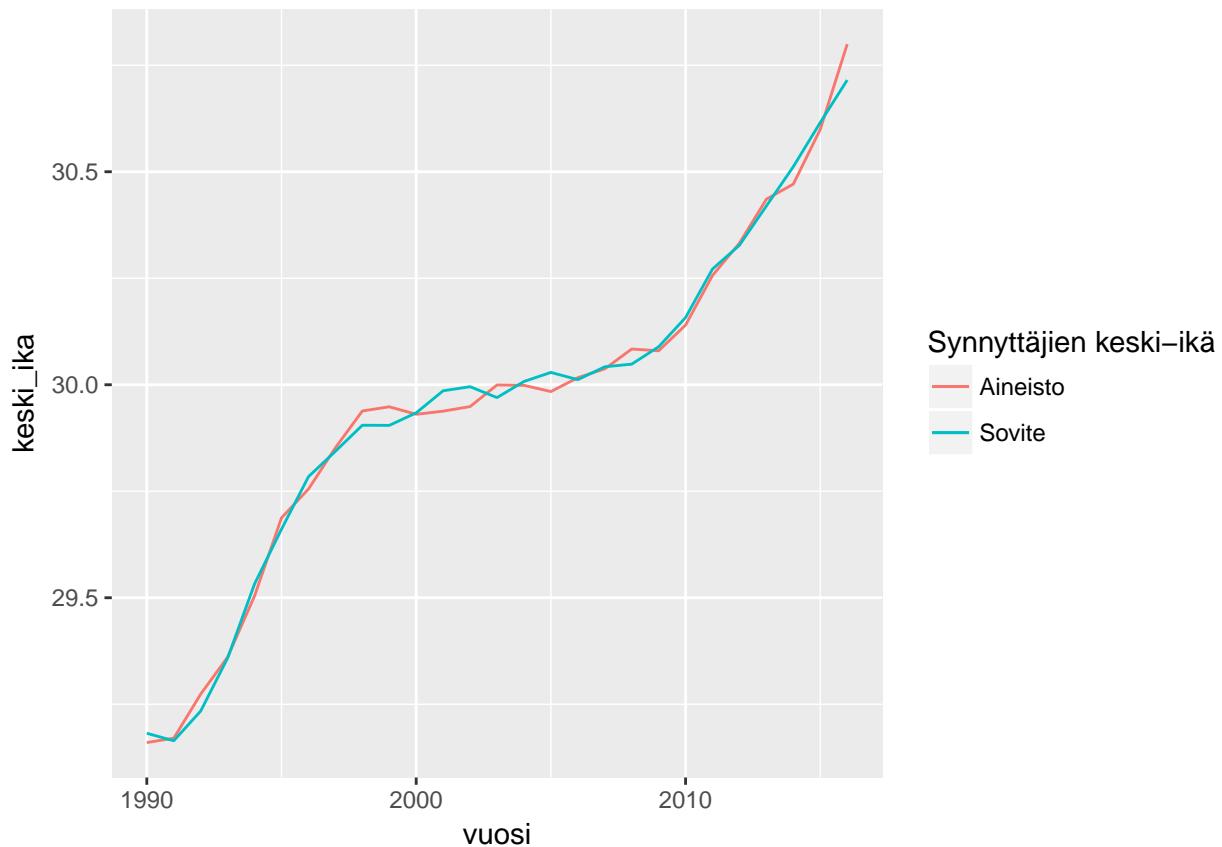
```

Regressiokerrointen mukaan 25-45 -vuotiaiden naisten työllisyysaste on tärkeä selittää synnyttäjien keski-iälle. Sen nousu alentaa keski-ikää. Muita tärkeitä tekijöitä ovat maaseudulla asuvien osuus ja sisemmässä kaupungissa asuvien osuus. Maaseudulla asuvien osuuden alenemine nostaa synnyttäjien keski-ikää. Myös sisemmässä kaupungissa asuvien osuuden *kasvu* nostaa synnyttäjien keski-ikää.

Tarkastetaan vielä, millaisia residuaaleja jää selittämättä mallilla. Kuvion mukaan residuaalit käyttäytyvät



Predictions under BMA
 hyvin.
`bas.lm(synnyttajien_keskiika ~ pienituloisuus_25_34 + lapsikuolleisuus + ty ...`
 Ja lopuksi verrataan mallin sovitetta



Työsuhteenvaikuttavuus

Kokeillaan sitten vaikuttaako työsuhdetyyppi syntyvyyteen. Aineistona käytetään tilastokeskuksen tietoja työsuhteenvaikuttavuudesta (määräaikainen vai jatkuva) ja siitä, onko se kokoaikainen vai osa-aikainen. Aineisto on vuosilta 2007-2016. Synnytysikäisten naisten lukumäärää pystytään ennustamaan melko hyvin, eikä siihen vaikuta juurikaan muut tekijät, kuten vaikkapa työllisyys tai taloudellinen tilanne. Jos sellaiset tekijät vaikuttavat syntyvyyteen, vaikutukseen pitäisi näkyä hedelmällisyyssluvussa.

```
syntyvyyys_reg = bas.lm(hedelmällisyyssluku ~ pienituloisuus_0_17 + pienituloisuus_18_24 + pienituloisuus_25_34 + pienituloisuus_35_49 + synnyttajien_keskiika + jatkuvakokoaika_miehet + jatkuvaosaaiaka_miehet + maaraaikakoko_miehet + maaraaikaosa_miehet + jatkuvakokoaika_naiset + jatkuvaosaaiaka_naiset + maaraaikakoko_naiset + maaraaikaosa_naiset + kaupungissaasuvat + maaseudulla + sisakaupungissa + BF + PostProbs + R2 + dim + logmarg)
```

```
## Warning in bas.lm(hedelmällisyyssluku ~ pienituloisuus_0_17 + pienituloisuus_18_24 + : dropping 1 rows due to missing data
```

Verrattuna aiempiin tuloksiin, näkyvät samat tekijät vaikuttavan voimakkaimmin hedelmällisyysslukuun. Työsuhteenvaikuttavuus ei näytä vaikuttavan. Tulokset vielä numeroina

	P(B != 0 Y)	model 1	model 2	model 3	model 4
## Intercept	1.000	1.000	1.000	1.000	1.000
## pienituloisuus_0_17	0.306	0.000	0.000	0.000	0.000
## pienituloisuus_18_24	0.147	0.000	0.000	0.000	0.000
## pienituloisuus_25_34	0.379	0.000	0.000	0.000	1.000
## pienituloisuus_35_49	0.156	0.000	0.000	0.000	0.000
## synnyttajien_keskiika	0.776	1.000	1.000	1.000	1.000
## jatkuvakokoaika_miehet	0.162	0.000	0.000	0.000	0.000
## jatkuvaosaaiaka_miehet	0.200	0.000	0.000	0.000	0.000
## maaraaikakoko_miehet	0.159	0.000	0.000	0.000	0.000
## maaraaikaosa_miehet	0.179	0.000	0.000	0.000	0.000
## jatkuvakokoaika_naiset	0.187	0.000	0.000	0.000	0.000
## jatkuvaosaaiaka_naiset	0.333	0.000	1.000	0.000	0.000
## maaraaikakoko_naiset	0.166	0.000	0.000	0.000	0.000
## maaraaikaosa_naiset	0.346	0.000	0.000	1.000	0.000
## kaupungissaasuvat	0.543	0.000	0.000	0.000	1.000
## maaseudulla	0.557	1.000	1.000	1.000	0.000
## sisakaupungissa	0.990	1.000	1.000	1.000	1.000
## BF	NA	1.000	0.347	0.328	0.245
## PostProbs	NA	0.045	0.016	0.015	0.011
## R2	NA	0.959	0.965	0.965	0.964
## dim	NA	4.000	5.000	5.000	5.000
## logmarg	NA	19.957	18.899	18.842	18.552
##		model 5			
## Intercept	1.000				
## pienituloisuus_0_17	0.000				
## pienituloisuus_18_24	0.000				
## pienituloisuus_25_34	0.000				
## pienituloisuus_35_49	0.000				
## synnyttajien_keskiika	1.000				
## jatkuvakokoaika_miehet	0.000				
## jatkuvaosaaiaka_miehet	0.000				
## maaraaikakoko_miehet	0.000				
## maaraaikaosa_miehet	0.000				
## jatkuvakokoaika_naiset	0.000				
## jatkuvaosaaiaka_naiset	0.000				
## maaraaikakoko_naiset	0.000				
## maaraaikaosa_naiset	0.000				
## kaupungissaasuvat	1.000				

```

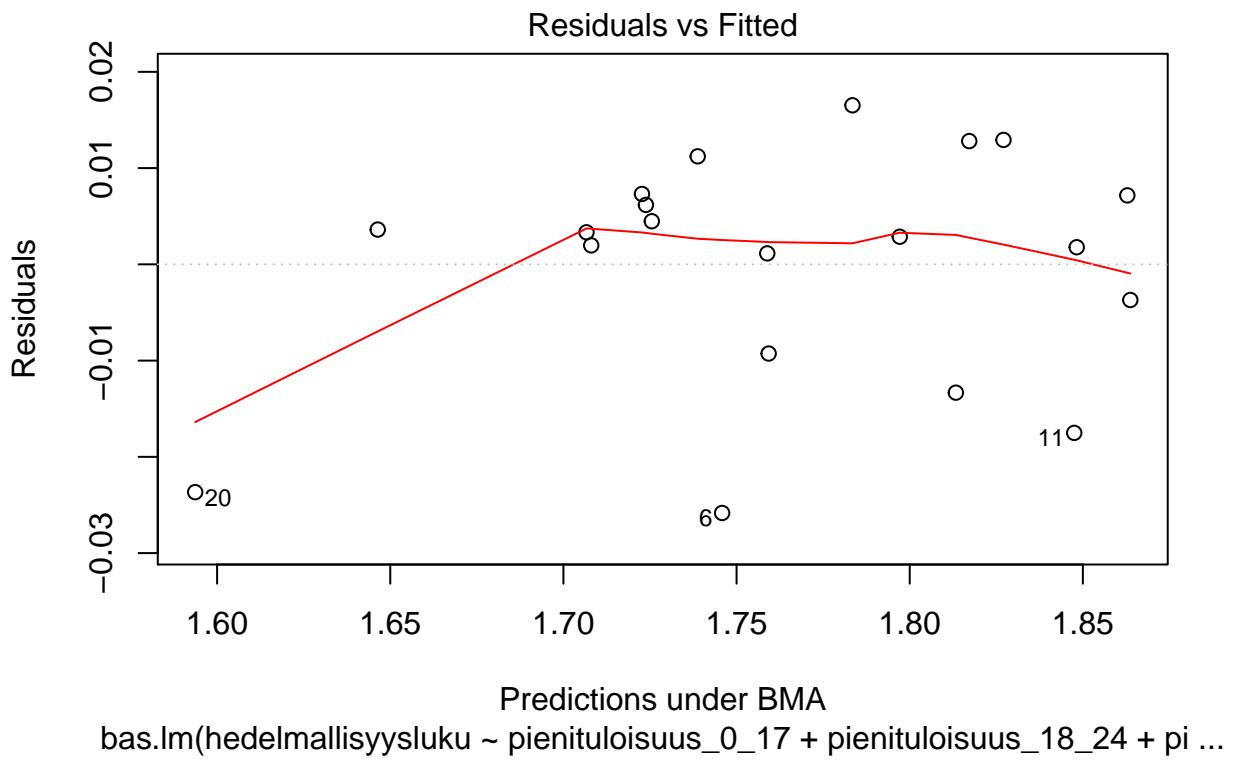
## maaseudulla          0.000
## sisakaupungissa      1.000
## BF                   0.240
## PostProbs            0.011
## R2                   0.951
## dim                  4.000
## logmarg              18.528

##
## Marginal Posterior Summaries of Coefficients:
##
## Using BMA
##
## Based on the top 65536 models
##                               post mean   post SD    post p(B != 0)
## Intercept                  1.764e+00  3.729e-03  1.000e+00
## pienituloloisuus_0_17     -2.358e-03  5.122e-03  3.058e-01
## pienituloloisuus_18_24    -9.147e-05  1.303e-03  1.472e-01
## pienituloloisuus_25_34    4.445e-03   7.588e-03  3.792e-01
## pienituloloisuus_35_49    9.675e-04   4.784e-03  1.564e-01
## synnyttajien_keskiika   -2.594e-01   1.791e-01  7.755e-01
## jatkuvakokoaika_miehet -6.561e-05   7.016e-03  1.619e-01
## jatkuvaosaaiaka_miehet  4.420e-03   1.506e-02  1.998e-01
## maaraaikakoko_miehet   -2.447e-04   7.788e-03  1.589e-01
## maaraaikaosa_miehet    -4.152e-03   2.195e-02  1.794e-01
## jatkuvakokoaika_naiset  6.656e-04   9.397e-03  1.874e-01
## jatkuvaosaaiaka_naiset  6.561e-03   1.336e-02  3.333e-01
## maaraaikakoko_naiset   -1.575e-03   1.237e-02  1.662e-01
## maaraaikaosa_naiset    -1.812e-02   3.177e-02  3.463e-01
## kaupungissaasuvat       5.474e-02   6.077e-02  5.434e-01
## maaseudulla             -5.344e-02   6.120e-02  5.572e-01
## sisakaupungissa         -2.594e-01   8.587e-02  9.899e-01

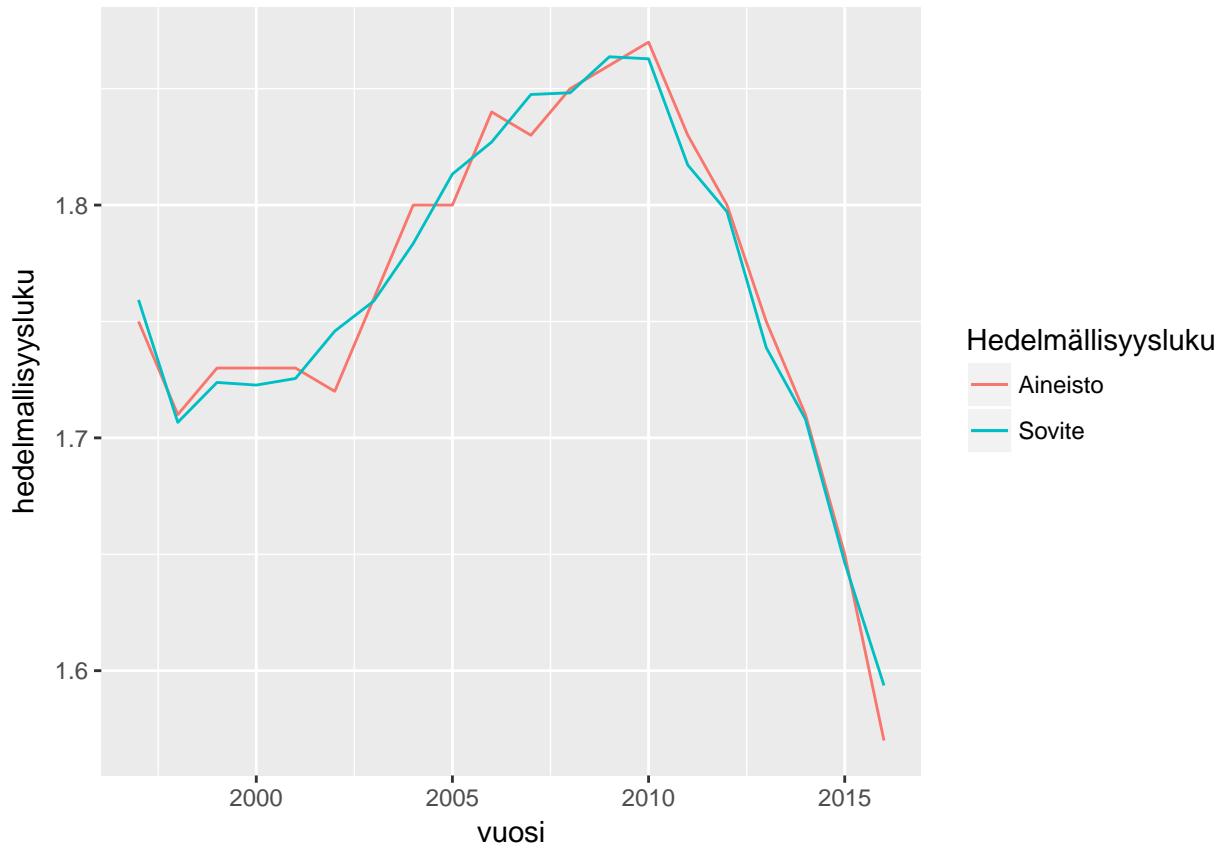
```

Kolmella faktorilla, sisemmässä kaupungissa ja maaseudulla asuvien osuudet sekä synnyttäjien keski-ikä, selittäävästi tässä aineistossa noin 96 % varianssista. Työsuhteiden tyypin merkitys on selvästi näitä vähäisempi. Regressiokerrointien etumerkit kertovat, että naisten jatkuviien osa- ja kokoaikaisten työsuhteiden osuuden kasvu lievästi kasvattaa hedelmällisyyslukua ja naisten määräaikaiset työsuhteiden osuuden kasvu saattaa hienoisesti alentaa hedelmällisyyslukua. Näillä tekijöillä ei kuitenkaan ole juurikaan selitysvoimaa, eivätkä ne selitä viime vuosien hedelmällisyysluvun laskua.

Tarkastetaan vielä, millaisia residuaaleja jää selittämättä mallilla. Numeroidut residuaalit ovat outliersita. Selvästi viime vuosien hedelmällisyysluvun lasku jää osin selittämättä myös tällä mallilla.



Ja lopuksi verrataan mallin sovitetta



Syntyvyyden laskun syyt

Vaikka tällä tutkimuksella ei pääse kiinni syntyvyyden laskun kausaalisiin syihin, näkyy tuloksista selvästi, että syntyvyyden laskua ajaa kaksi erilaista tekijää: hedelmällisyysluvun lasku ja synnytsikäisten naisten lukumäärän lasku. Vaikempaa on yrittää selittää, mistä hedelmällisyysluvun lasku johtuu.

Yksi tässäkin tutkimuksessa selvästi esille tullut tekijä hedelmällisyysluvun laskulle on synnyttäneiden keski-iän nousu. Se kertoo selvästi siitä, että 20-24 -vuotiaat naiset saavat lapsia selvästi aiempaa vähemmän, samoin 25-29 -vuotiaat. Tämä saattaa johtua asennemuutoksesta, mutta myös kaupungistuminen, pienituloisuus ja työttömyys kytkeytyivät tässä tutkimuksessa synnyttäneiden keski-iän nousuun. Synnyttäneiden keski-iän noussee, koska 20-24 -vuotiaat saavat vähemmän lapsia, mutta vahvin kytkentä keski-iän nousuun on kaupungistumisella ja 25-34 -vuotiaiden naisten työllisyysasteella.

Tilastokeskuksen tiedoista tiedetään, että kaupungeissa syntyvyys on alempaa kuin muussa maassa: "Vastaavasti matalin syntyvyys oli jaksolla 2013–2017 Helsingissä, jossa kokonaishedelmällisyysluku oli 1,27. Turussa luku oli 1,29 ja Tampereella 1,34. Myös kaikkien kuntien vertailussa viimeksi mainittujen kaupunkien syntyvyys oli matalimpien joukossa." (Tilastokeskus, väestörakenne 2018). On kuitenkin epäselvä, missä määrin kaupungistuminen aiheuttaa alempaa syntyvyyttä. Havainnot urbanisaation assosiaatiosta alempaan syntyvyyteen tukee tätä. Regressiokerrointen mukaan sisemmassä kaupungissa asuvien osuuden kasvu alentaa syntyvyyttä ja lykkää lasten hankintaa. Toisaalta ulommassa kaupungissa (samoin kuin maaseudulla) asuvien osuuden kasvu näyttäisi lisäävän syntyvyyttä ja vähentävän lastenhankinnan lykkäämistä. Tulokset ovat intuitiivisesti hyväksyttävissä, koska sisemmissä kaupungeissa erityisesti lapsiperheille sopivien asuntojen hintataso on korkeampi, kun taas ulompana hintataso on alempi. Kaupungistuminen ei siis sinänsä alenna syntyvyyttä, vaikka sisäkaupungeissa asuvien osuuden kasvu näyttää alentavan syntyvyyttä.

Hiilamo (2017) arvioi, että työttömyys kytkeytyy hedelmällisyyslukuun merkityksellisellä regressiokertoimella -0,12. Se on merkittävästi suurempi kuin tässä tutkimuksessa löytyneet regressiokertoimet (suuruusluokka -0,01), jotka eivät olleet kovin vahvasti kytkeytyneitä. Työllisyysasteet tuottivat hieman parempi tulokset kuin työttömyysasteet.

Pienituloisuus kytkeytyi tutkimuksessa lievästi hedelmällisyysluvun laskuun. Eurostatin mukaan materiaalinen deprivatio (Severe material deprivation) on Suomessa vähentynyt yhtäjaksoisesti vuodesta 2010 alkaen.

Avioituvuus erityisesti 30-34-vuotiailla kytkeytyy selvästi syntyneiden lasten määrään. Suurin osa lapsiperheistä on edelleen avioperheitä. Ainakin aiemmin on argumentoitu, että verrattuna avoliittoihin, avioliitoissa syntyy enemmän lapsia. Kyse voi olla myös valinnasta, mutta osaltaan avioituvuuden aleneminen kertoo asennemuutoksesta.

Tässä tutkimuksessa aineisto koostui pienestä määrästä havaintoja. Tästä syystä tuloksiin on syytä suhtautua varovaisuudella, vaikka usealle löydetylle tekijälle on selkeä selitys.

Johtopäätökset

Syntyvyyden laskua selittää muutokset synnytsikäisten naisten lukumäärässä ja hedelmällisyysluvussa. Synnytsikäisten naisten lukumäärä muuttuu hitaasti, mutta hedelmällisyysluku on laskenut nopeasti vuoden 2010 jälkeen. Hedelmällisyysluvun muutosten selittäminen onnistuu tarkastellulla malleilla vaillimaisesti, mutta avioituvuudella ja kaupungistumisella vaikuttaisi olevan kytkentä hedelmällisyysluvun laskuun. Sitä vastoin tarkastellussa aineistossa eläkemenolla, alle 1-vuotiaiden kuolleisuuden laskulla tai työsuhdetyyppien muutoksella ei näytä olevan juurikaan kytkentää hedelmällisyyslukuun.

Viittaukset

Becker, G.S., Barro, R.J., Reformulation of the economic theory of fertility, The quarterly journal of economics, 1988

Boldrin, M., De Nardi, M., Jones, L.E. Fertility and social security, Journal of Demographic Economics 81, 261-299, <https://doi.org/10.1017/dem.2014.14>, 2015

Doepke, M., Child mortality and fertility decline: Does the Barro-Becker model fit the facts?

Hilamo, H. T., Fertility Response to Economic Recessions in Finland 1991–2015 Finnish Yearbook of Population Research 52, 15-28 . DOI: 10.23979/fypr.65254, 2017

Kohler, H.-P., Kohler, I., Fertility Decline in Russia in the Early and Mid 1990s: The Role of Economic Uncertainty and Labour Market Crises, European Journal of Population 18, 233-262, <https://doi.org/10.1023/A:1019701812709>, 2002.

Martine, G., Alves, J.E., Cavenaghi, S., Urbanization and fertility decline: Cashing in on structural change, IIED Working paper, 2013

Miettinen, A., Miksi syntyyvyyks laskee? Suomalaisten lastensaantiin liittyviä toiveita ja odotuksia. Perhebarometri 2015. Väestöliitto, 2015.

Miettinen, A., Jalovaara, M. Stable employment – more babies? Life stage and educational differences in the effects of labour market attachment on first birth among Finnish men and women. Working Papers on Social and Economic Issues 15/2016, Turku Center for Welfare Research, 2016

Tikanmäki, H., Huomisen aikuiset syntyyvät nyt, Eläketurvakeskuksen blogi <https://www.etk.fi/blogit/huomisen-aikuiset-syntyyvat-nyt/>, 2017

Liite: Aika-sarjojen kuvaukset

syntyneet Syntyneiden lasten lukumäärä. Lähde: Tilastokeskus

hedelmällisyysluku "Kokonaishedelmällisyysluku saadaan laskemalla yhteen yhdelle ikävuodelle lasketut hedelmällisyysluvut. Näin saatu luku tarkoittaa laskennallista lasten määrää, jonka nainen synnyttää kyseisen vuoden hedelmällisyyden pysyessä voimassa naisen koko hedelmällisen kauden edellyttäen, ettei hän kuole ennen tämän kauden päättymistä." Lähde: Tilastokeskus, Väestörakenne-tilasto

pienituloisuus_0_17, pienituloisuus_18_24, pienituloisuus_25_34, pienituloisuus_35_49, pienituloisuus_50_64 "Pienituloisuusaste: Pienituloisten osuus kyseisessä väestöryhmässä. Pienituloisia henkilötä: Pienituloisia henkilötä kyseisessä väestöryhmässä. Pienituloisten mediaanitulo: Pienituloisten ekvivalentit mediaanitulot kyseisessä väestöryhmässä. Köyhyysvaje: Pienituloisten tulomediaanin ja koko väestön pienituloisuusrajan erotus prosentteina rajasta. Ekvivalenttitulo: Kotitalouden käytettävissä olevat rahatulot kulutusyksikköä kohden vuodessa. Kulutusyksikkö: Modifioitu OECD-asteikko (1.0 - 0.5 - 0.3). Pienituloisuusrajet lasketaan vuosittain kaikkien kotitalouksien ekvivalenttitulon mediaanin perusteella. Julkaisuissa käytetty virallinen pienituloisuusraja on 60 % mediaanitulosta ja tulokásite käytettävissä olevat rahatulot. Euromääritiset tiedot on esitetty vuoden 2016 rahassa." Lähde: Tilastokeskus, Tulonjakotilasto

lapsikuolleisuus Alle 1-vuotiaana kuolleiden osuus elävänä syntyneistä. Lähde: Human mortality database

työllisyysaste_miehet_15_24_lag1, työllisyysaste_naiset_25_34_lag1, työllisyysaste_miehet_35_44_lag1, työllisyysaste_naiset_15_24_lag1, työllisyysaste_miehet_25_34_lag1, työllisyysaste_naiset_45_44_lag1 Yhden vuoden viivästetty 10-vuotisikäuokittainen työllisyysaste miehille ja naisille. Lähde: Tilastokeskus

synnyttajien_keskiika Laskettu synnyttäneiden äitien ikien keskiarvona. Lähde: Tilastokeskus

avioliitot Lähde: Tilastokeskus

elakemeno Kokonaiseläkemeno Suomessa Lähde: Eläketurvakeskus

kaupungissaasuvat, maaseudulla, sisakaupungissa Kaupungeissaasuvien osuus väestöstä, maaseudulla asuvien osuus väestöstä ja Lähde: Tilastokeskus

jatkuvakokoika_miehet, jatkuvaosaika_miehet, maaraaikakoko_miehet, maaraaikaosa_miehet, jatkuvakokoika_naiset, jatkuvaosaika_naiset, maaraaikakoko_naiset, maaraaikaosa_naiset Lähde: Tilastokeskus, työvoimatutkimus

naisten_lkm_15_19, naisten_lkm_20_24, naisten_lkm_25_29, naisten_lkm_30_34, naisten_lkm_35_39, naisten_lkm_40_45, naisten_lkm_45_49 Synnytyskäiset naiset 5-vuotisikäluokittain Lähde: Tilastokeskus

miesten_tyottomyysaste_15_24_lag1, naisten_tyottomyysaste_15_24_lag1, miesten_tyottomyysaste_25_34_lag1, naisten_tyottomyysaste_25_34_lag1, miesten_tyottomyysaste_35_44_lag1, naisten_tyottomyysaste_35_44_lag1 Yhden vuoden viivästetty 10-vuotisikäluokittainen työttömyysaste miehille ja naisille Lähde: Tilastokeskus

Liite: Kaupunkialueet

Sisempi kaupunkialue Sisempi kaupunkialue kuvaavat kaupunkien tiivistä ja yhtenäistä tehokkaasti rakennettua aluetta. Sisemmällä kaupunkialueella lähiympäristön vallitseva aluetehokkuus on vähintään 0,1, joka vastaa tiivistä pientaloaluetta. Sisemmästä kaupunkialueesta irrallaan sijaitsevat korkean aluetehokkuuden alueet sisällytetään ulompaan kaupunkialueeseen. Jokaisella keskustaajamalla voi olla vain yksi sisempi kaupunkialue. Poikkeuksena tätä muodostavat Helsingin kaupunkiseudulla sijaitsevat kaupungit Järvenpää ja Kerava, joille muodostetaan oma sisempi kaupunkialue. Lähde: Ympäristöministeriö

Ulompi kaupunkialue Ulompi kaupunkialue koostuu erillisistä lähiöistä, kaupan, teollisuuden ja toimistojen reunakaupungista sekä toisaalta kaupungin viheralueista ja alemman tehokkuuden asuinalueista, jotka kuitenkin ovat selvästi osa yhtenäistä kaupunkialuetta. Ulommalla kaupunkialueella kolmasosa lähiympäristöstä on vähintään 0,02 aluetehokkuuden aluetta. Tällä kriteerillä ulomman kaupunkialueen ulkoraja estimoi asemakaavoitetun alueen ulkorajaa ja näin suunnitellusti rakennettua yhtenäistä kaupunkialuetta. Lähde: Ympäristöministeriö