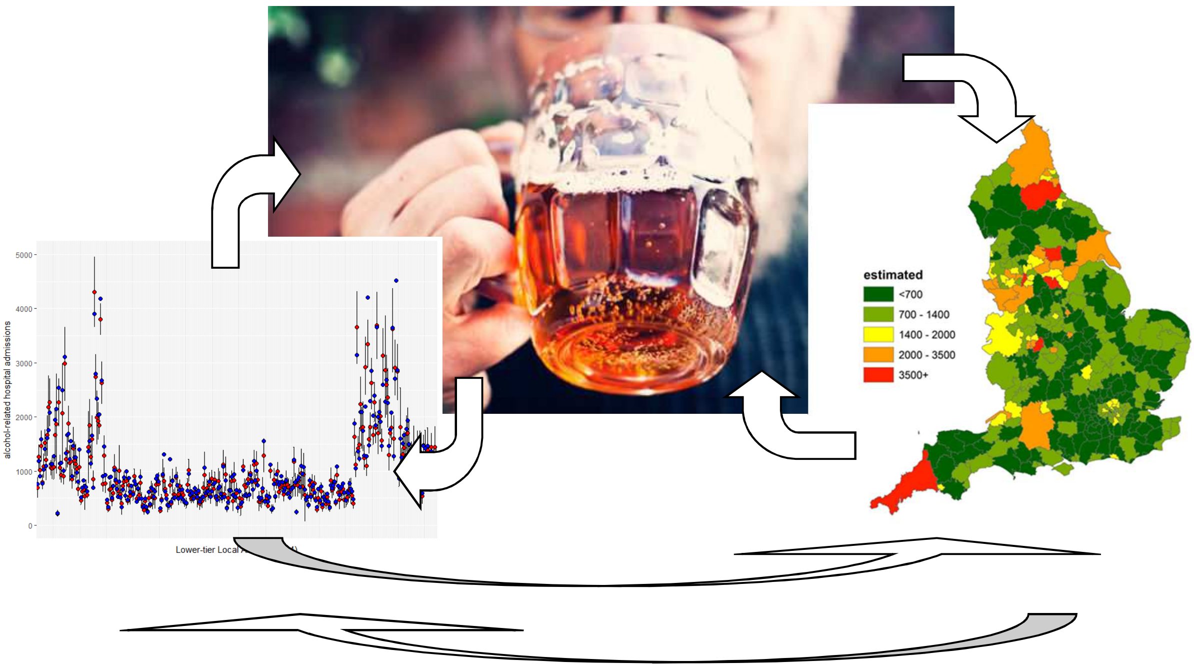
‘Nowcasting' of population alcohol-related harms using novel Bayesian timeseries methods and synthetic controls.

SG 16/17 235

Dr Frank de Vocht, Ms Cheryl McQuire, Professor Kate Tilling, Professor Matthew Hickman



Alcohol Research UK Final Report

ACKNOWLEDGEMENTS
The authors would like to thank Peter Levell from the Institute for Fiscal Studies for providing the original data on time series of real alcohol duties used in Chapter 'Example: impact of alcohol taxation' of this report.

AUTHOR DETAILS
Dr Frank de Vocht. Senior Lecturer in Epidemiology and Public Health Research.
Population Health Sciences, Bristol Medical School
University of Bristol, Bristol, UK. Canynge Hall, 39 Whatley Road, Bristol, BS8 2PS.

Email: frank.devocht@bristol.ac.uk.
Tel: +44.(0)117.928.7239
Postal address: Canynge Hall, 39 Whatley Road, Bristol, BS8 2PS.

Ms Cheryl McQuire, Research Associate in Epidemiology and Alcohol Research
Population Health Sciences, Bristol Medical School
University of Bristol, Bristol, UK.
Email: cheryl.mcquire@bristol.ac.uk

Professor Kate Tilling. Professor of Medical Statistics.
Population Health Sciences, Bristol Medical School
University of Bristol, Bristol, UK.
Email: kate.tilling@bristol.ac.uk

Professor Matthew Hickman. Professor of Public Health and Epidemiology.
Population Health Sciences, Bristol Medical School
University of Bristol, Bristol, UK.
Email: matthew.hickman@bristol.ac.uk

This report was funded by Alcohol Research UK. Alcohol Research UK is an independent charity working to reduce alcohol-related harm through ensuring policy and practice can be developed on the basis of reliable, research-based evidence.
www.alcoholresearchuk.org

CONTENTS

EXECUTIVE SUMMARY

Fore- and nowcasting of expected population trends in the burden of alcohol-related harms is undertaken to help plan for the future. It would, therefore, be beneficial to have accurate forecasts that can inform policy making. Forecasts would ideally be available at different geographical levels to provide additional information on local variations in the expected burden of alcohol, or the likely effect of specific policies in different areas of the country. Although a plethora of forecasting methods exist, they often rely on straightforward linear projections, or on a static set of estimated associations between time-series of interest and predictors.

This project is a 'proof of principle' study in which a novel statistical methodology, Bayesian structural timeseries, is used to model 'known' time-series of 2002/3 to 2015/6 alcohol-related hospital admissions for all lower-tier local areas in England (n=326), and to forecast projections 5 years into the future (i.e. to 2021), based on projected local trends in demographics.

This methodology may be more beneficial than other methods as it allows for temporal trends in the associations between outcomes and predictors to be incorporated, and Bayesian model averaging makes forecasting minimally dependent on specific hypothesized (and therefore potentially incorrect) model specifications.

Prior to utilizing the methodology for forecasting purposes, we investigated prediction errors from nowcasting to evaluate the accuracy of the models and to provide information on the quality of the forecasts. In addition, the novel methodology is further used to estimate the impact of forecasted trends in alcohol taxation on 2021 alcohol related hospital admissions in England, and to explore the geographical variations therein.

The Bayesian structural time-series approach generates relatively accurate and acceptable estimates of time trends with an average 5-year forecasting error of approximately -2% and average local error 13%. At the national level, the annual number of alcohol-related hospital admissions if no other changes occur is forecasted to increase by an extra 4,265 by 2021. However, because the population size is forecasted to increase more, this in relative terms corresponds to a small reduction in the crude rate (per 1,000 people) of 2.3%.

Trends in hospital admissions differ between regions, local areas, and types of local areas. They are expected to increase most (relative to 2016) in the North West and least in the London Cosmopolitan areas. Aggregation by type of lower tier local authority (LTLA) suggests the largest increases are forecasted in the deprived, but also the prosperous, semi-rural and seaside living areas and largest decreases in metropolitan and cosmopolitan areas and in University towns and cities. These trends seem to be related to alcohol harms disproportionally affecting more deprived communities combined with a demographic secular trend of middle-aged and older adults moving away from cities to (semi-) rural and seaside towns.

The 5-year forecasted impact from forecasted real alcohol taxation on alcohol-related hospital admissions is, overall, small but beneficial, with a 0.2% decrease in expected alcohol-related hospital admissions. Alcohol taxation is expected to lead to a greater reduction in alcohol-related hospital admissions in areas with higher levels of deprivation.

This 'proof-of-principle' study showed that this novel modelling framework provided accurate forecasts of secular trends in local alcohol-related hospital admissions in England, and may well be useful for future policy making at the local, regional and national level. Future work will include detailed assessment of the precision of these estimates, optimization of model specifications including the possibility of including additional modifiable factors, and will also explore applications to other scenarios, including non-alcohol related outcomes.

BACKGROUND

Fore- and nowcasting of expected population trends in the burden of alcohol-related harms is undertaken in order to plan for the future and to examine whether policies are likely to lead to positive or negative impacts. However, current methods are mainly based on straightforward extrapolation of observed trends and standard regression models. For example, the Office for National Statistics (ONS) uses ARIMA (autoregressive integrated moving average) models for nowcasting of healthcare output (ONS, 2012), and is currently actively consulting on new methods (ONS, 2016). Specifically related to alcohol-related hospital admissions, the Local Alcohol Profiles for England (LAPE)have developed a forecasting tool (the 'National Planning Tool') (http://webarchive.nationalarchives.gov.uk/20171107173502/http://www.lape.org.uk/data.html), which also is based on straightforward regression and rather rudimentary. In other policy areas however, development of fore- and nowcasting methodologies has received more attention, and sophisticated nowcasting methods are employed in for example the atmospheric sciences and meteorology (Met Office, 2011), marketing and (macro)economics (Giannone et al. 2008) and, of particular interest within alcohol research, within health sciences to predict influenza-like illness rates (Lampos et al. 2015) and Dengue epidemics (de Almeida Marques-Toledo et al. 2017), for example.

Improvement of nowcasting methodologies in public health, and alcohol research therein, would be beneficial and could have important implications for prioritization, development and implementation of policies, including licensing decisions and screening and brief interventions. Therefore, this project aimed to utilize a novel analysis framework, originally developed by Google, to evaluate the impact of their marketing campaigns (Scott and Varian. 2013) to improve now/forecasting for alcohol research. We previously used this methodology within a counterfactual causal inference framework to estimate the impact of local alcohol licensing policies on local alcohol-related hospital admissions and crime rates using a natural experiment study design (de Vocht et al., 2017a). Here, the statistical methodology – Bayesian structural timeseries – will be used independently from the causal inference framework to project forward expected trends in alcohol-related hospital admissions for each of 326 local authorities in England.

> ***Forecasting*** is the process of making predictions of the future based on past and present data and most commonly by analysis of trends [https://en.wikipedia.org/wiki/Forecasting].
>
> ***Nowcasting*** is the prediction of the present, the very near future and the very recent past [Bańbura et al. 2010; https://en.wikipedia.org/wiki/Nowcasting_(economics)]

This novel methodology was developed and used mainly in management and (macro)economics, but is very promising in the context of public health. Alcohol research would greatly benefit from improved fore- and nowcasting capabilities at national and local level, and this study is a proof-of-concept study for further development.

This fits in with the goals of Alcohol Research UK that to find 'the most effective ways of improving policy and practice to tackle them' and that 'alcohol policy and practice should be informed by the best available research', while moreover it directly addresses priority area four 'Alcohol research methods. – how can we enhance and refine the many ways in which research is carried out across the alcohol field?' and contributes to policy (priority two) and education (priority 5).

Practically, the project has provided improved now- and forecasting of alcohol-related harms at the local and national level. This could help guide the design and implementation of licensing or other alcohol policies, and/or better targeting of screening and brief interventions. The methodology is flexible, thereby facilitating continuous (or periodical) updates of these now/forecastings when new data become available.

PROJECT AIM

The original aim of the project was to investigate the use of a novel statistical method – Bayesian structural time series – to improve predictions (now- and forecasting) of alcohol-related harms and crime rates at national and at local (lower-tier local authority; LTLA) level.
However, parallel to the writing of this research grant application, the investigators conducted work on local-level alcohol-related crime rates in England (de Vocht et al. 2017a,b) in which it became apparent that there are problems with the reporting of crimes that make the evaluation of time series difficult. It was felt that this would complicate the evaluation of a new method in a new context. Therefore alcohol-related crime rates were not included in the current project.
It was further noted that freely available and reliable forecasting data that would be informative in relation to alcohol-related hospital admissions, especially at the local level, was more limited than expected. Nonetheless, time series of alcohol taxation that included expected future trends were obtained at the national level and have been included as a case study. Furthermore, forecasting information was available for the age distribution in England, which has also been made available at local authority level. The decision was made therefore to modify the project aim to include this local variation. This is important because it directly addresses the issue of where in England future health harms are expected to occur, as a result of demographic changes alone.

As such, the ***modified project aim*** is:

To investigate the use of a novel statistical method – Bayesian structural time series – to improve predictions (now- and forecasting) of alcohol-related harms at the local level as a result of demographic changes, and to explore the potential of alcohol taxation on these predictions.

PROJECT LAYOUT

The project development and evaluation phase consisted of three distinct phases:

I. ***Nowcasting the present (development and evaluation)***. Using the available and linked data for the years 2002/3 to 2015/16 (described below), a Bayesian structural time series model was developed for the time series of alcohol-related hospital admissions, where these were known. This enabled comparison of the forecasts to the actual lower-tier local area measured alcohol-related hospital admissions and subsequently assessment of accuracy of the forecasting models. Furthermore, using various performance statistics the optimal model specification was established.

II. ***Forecasting the future (implementation)***. Once it was demonstrated that the Bayesian structured time series provided acceptable now- and forecasts, these local models were run 5 years into the future (i.e. to 2021) to forecast the expected trends in annual alcohol-related hospital admissions and their rates (per 1,000 people), based on trends in population sizes and age distribution. Aggregation of the LTLA local models in larger geographical or area-type categories enabled more generic inferences about expected health burdens across England.

III. ***Assessing the impact of alcohol taxation (example)***. As an example of how the new forecasting framework can be used to forecast the impact of policies, real alcohol duty rates (in real terms, relative to 1978) for beer, wine and spirits, and forecasted to 2021 (The Institute for Fiscal Studies' (IFS) Green Budget 2016) were obtained and included in the demographics models developed in (II).

METHODS
Data
Annual alcohol-related hospital admission counts for all 326 lower-tier local authority (LTLA) areas were obtained from the Local Alcohol Profiles for England (LAPE) website for the midterm years 2002/3-2015/6. These were linked to mid-year population figures for persons, stratified in 5-year age groups, for local authorities in England, which were downloaded from The National Archives [1], and to forward projections from 2014 up to 2039 [2]. For modelling purposes, these were subsequently aggregated to time-series of 10-year strata: 0-15 years of age, 16-24, 24-34, 35-44, 45-54-55-64, 65-74, and 75+ years of age.

[1] http://webarchive.nationalarchives.gov.uk/20160105223339/http://www.ons.gov.uk/ons/rel/pop-estimate/population-estimates-for-uk--england-and-wales--scotland-and-northern-ireland/index.html

[2] https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/population projections

Results have been presented in Tables and in geospatial maps. For the latter, results were exported into ArcGIS (version 10) and linked to Local Authority Districts Boundaries for England (December 2015) downloaded from:

http://geoportal.statistics.gov.uk/datasets/local-authority-districts-december-2015-full-clipped-boundaries-in-great-britain.

For all 326 lower-tier local authorities, and spatial maps designed accordingly.

Statistical methodology
Bayesian structural time-series models [https://en.wikipedia.org/wiki/Bayesian_structural_time_series] use a Bayesian state-space time-series model that includes one component of state as a linear regression on the predictors. The methodology is developed as an R package (the *bsts* package [https://cran.r-project.org/web/packages/bsts/bsts.pdf]), which was specifically developed for this purpose) by Scott and Varian and described in detail in by Scott and Varian (2014) and Broderson et al (2015).

In summary, the outcome is represented as the sum of three components; a time series component that exploits the temporal patterns in the data; a regression component that relates the outcome to the covariates; and an error component that accounts for any unexplained variability. Two linked equations describe how the outcome data, i.e. the alcohol-related hospital admissions, (y_t;) are linked to an unobserved latent state (α_t) (the observational equation; Eq. 1), and how the latter changes over time (the 'transition equation; Eq. 2):

$$y_t = Z_t^T \alpha_t + \varepsilon_t \quad \text{(Eq. 1)}$$
$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t \quad \text{(Eq. 2)}$$

where $\varepsilon_t \sim \mathrm{N}(0, \sigma_t^2)$ and $\eta_t \sim N(0, Q_t)$, and are both independent of all other unknowns, y_t is a scalar observation, Z_t is an output vector, T_t is the transition matrix, R_t the control matrix, η_t describes the system error with a state-diffusion matrix Q_t ($R_t \eta_t$ allows the inclusion of seasonality in these analyses, which is not used here), and ε_t is a scalar observation error with variance σ_t. The errors of different state-component models are assumed to be independent.

State space models (SSMs) refer to a class of probabilistic graphical models that describe the probabilistic dependence between the latent state variable and the observed measurement. The state or the measurement can be either continuous or discrete. The term "state space" originated in 1960s in the area of control engineering). SSM provides a general framework for analyzing deterministic and stochastic dynamical systems that are measured or observed through a stochastic process. The SSM framework has been successfully applied in engineering, statistics, computer science and economics to solve a broad range of dynamical systems problems. Other terms used to describe SSMs are hidden Markov models (HMMs) and latent process models. The most well studied SSM is the Kalman filter, which defines an optimal algorithm for inferring linear Gaussian systems [Chen and Brown, 2013; http://www.scholarpedia.org /article/State_space_model]

The model further includes a regression component which enables the construction of a synthetic time series based on combinations of the covariates, which uses a generalisation of the local linear trend model with the slope having a stationary random walk (described by equations 3 and 4):

$$\mu_{t+1} = \mu_t + \delta_t + \eta_{\mu,t} \quad \text{(Eq. 3)}$$
$$\delta_{t+1} = D + \rho(\delta_t - D) + \eta_{\gamma,t} \quad \text{(Eq. 4)}$$

Where μ_t is the value of the trend at time t and δ_t the slope at time t, $\eta_{\mu,t} \sim N(0, \sigma_\mu^2)$, and $\eta_{\delta,t} \sim N(0, \sigma_\delta^2)$. The two components of η are independent and *D* signifies the long-term slope.

A spike-and-slab prior is placed on the regression coefficients and the model averages over all covariates are then used to construct the outcome time series from the covariate time series. The "spike" determines the probability that a covariate has a non-zero coefficient (i.e. being included in the model) based on independent Bernoulli distributions, and the "slab" is a weakly informative Gaussian prior with a large variance.

The posterior predictive density is a joint distribution over all data points rather than a collection of pointwise univariate distributions to ensure a correct serial structure (Scott and Varian, 2014).
Prior distributions for the variance are set as Gamma distributions, and for the local linear trend models the incremental error in the state is *a priori* assumed to be small and set to $\frac{1}{\sigma^2} \sim G(10^{-2}, 10^{-2} s_y^2)$, where $s_y^2 = \sum_t (y_t - \bar{y})/(n-1)$.

Posterior simulation is done using a Gibbs sampler and Kalman filter o simulate from a Markov chain with a stationary distribution (for model parameters θ) *p(*θ*, a* | $Y_{1:n}$*)*, and the posterior incremental effect is described by p($\tilde{Y}_{n+1:m}$ | $Y_{1:n}$, $X_{1:m}$), where $\tilde{y}_{n+1}, \ldots, \tilde{y}_m$. The posterior distribution of the time series can then be projected forward based on the known time-series data and the weighted projected time series.

For the assessment of prediction errors The *CausalImpact* package (also for *R*) is used (Broderson and Hauser, 2017), which uses the *bsts* package as described above, but automates the comparison of the projected and measured time series.

In contrast to what was mentioned in the research proposal, dynamic time warping matching is not required for the current analyses.

Model Set-up

The outcome data were linked to the predictor data for all 326 local authorities. City of London and Isles of Scilly were dropped because of missing data and small population sizes. As a result, the analyses in this report are based on 324 areas. The impact of the two missing LTLAs on nationally aggregated alcohol-related hospital admissions is minimal.
An *R* script was developed to automatically create individual datasets for each of the LTLAs. These consisted ofthe time series for the local area alcohol-related hospital admissions and the 9 time-series of the age groups, and subsequently run the LTLA-specific Bayesian structural time series models, extract the relevant parameters, and combine these in a new outcome dataset for further processing.

The following Bayesian priors were set:

- Model parameters (i.e. β) expected to be non-Null in each model = 5
- Expected explained variance (R^2) = 70%. Based on preliminary non-Bayesian run.
- Degrees of freedom set to length of 'known data' minus 1 = 12 (nowcasting models; i.e. 2003 to 2016 minus 1), with different values for evaluation analyses; depending on analysis.
- μ (i.e. the estimate of alcohol-related hospital admissions) set to a normal distribution with:
  - μ = arithmetic mean of 'known data' alcohol-related hospital admission time series.
  - Initial value of mu = first value of alcohol-related hospital admissions in time series.
  - Variance of μ = 1% of standard deviation (σ) 0 of 'known data' of alcohol-related hospital admissions; set because these are relatively stable time series.
- σ for μ set to inverse Gamma distribution with:
  - Initial value of σ = 5% standard deviation of 'known data' of alcohol-related hospital admissions.

- Upper limit σ = 150% standard deviation of 'known data' of alcohol-related hospital admissions.
- Sample size = prior observation count 'known data'; depending on model (13 for nowcasting analyses).

Because predictions are a trade-off between 'model freedom' and 'statistical power', priors for standard deviations of variance were based on preliminary evaluations and optimizations of prediction errors. Note that that model predictions are relatively insensitive to these choices, but the precision of the estimates is affected.

In initial runs of individual LTLA models the Markov Chain Monte Carlo (MCMC) chains were evaluated based on the following criteria:

- Visually by trace and density plots
- *Raftery-Lewis* diagnostic tests with default accuracy of 0.005 to evaluate mixing, correlation and inappropriate starting values (interpreting Dependence factor (I) > 5 as indicative of problems).
- *Geweke* diagnostics and *Heidelberger-Welch* tests to evaluate MCMC chain stability.
- Mean and range of one-step prediction errors (i.e. a measure of the deviation of the estimation from real, measured outcomes) were calculated and *Durbin-Watson* tests used to evaluate residual correlation in these [Cowles and Carlin, 1996].

Based on these initial exploratory analyses, to ensure all evaluation criteria were met and all estimations were based on stable distributions, estimation was based on 20,0000 MCMC samples following a 10% burn-in period.
Posterior tail-areas probabilities, or posterior predictive p-values, were calculated, and interpreted as the posterior mean of classical p-values [Meng, 1994], using 5% cut-offs to evaluate precision.

This report will focus primarily on the accuracy of the predictions and how the methodology can be utilised. The precision of these estimates would also be important for informing policy decisions, however, this is highly dependent on the model specifications and Bayesian priors (while the accuracy is to a much lesser extent), and requires additional work that is beyond the scope of this report.

RESULTS

Assessment of nowcasting error

Evaluation of prediction errors were done using the first 10 years of data (2003-2013) to predict alcohol-related hospital admissions forward for 1, 2, and 3 years; 3 years nowcasting being the maximum where measured data are available for the full year, thereby enabling assessment of prediction errors. In addition, 5-year nowcasts are calculated, but these are based on 8 years of observed data only. A summary of the results is found in Table 1:

Table 1. Summary of nowcasting results.

	Nowcasting period			
	1 year	**2 years**	**3 years**	**5 years**
National-level alcohol related hospital admissions				
measured	326,940	326,950	336,330	336,330
Nowcasted(*)	325,971	327,080	327,952	329,061
Accuracy	99.7%	100.0%	97.5%	97.8%
LA-level alcohol related hospital admissions				
Average of all LTLAs	1,009.1	1,009.0	1,038.1	1,038.1
Absolute average error	64.3	82.7	107.9	138.4
% error	6.4%	8.2%	10.4%	13.3%
MSE**	8693.9	15,959.9	25,589.9	38,000.8

** nowcasted for 324 LAs individually and then aggregated up to national level.*
*** Mean-squared error (MSE)*

As shown, the accuracy in the nowcasts is good and, after aggregating at national level, ranges from 97.5% to 100% in different years. At the local level, accuracy of individual nowcasts is good, but decreases over the nowcasting period from 6.4% in year 1 to 13.3% for year 5. In absolute terms, the average absolute nowcasting error increases from 64 alcohol-related hospital admissions in year 1 to 138 in year 5. When interpreting these figures, however, it is also important to realise that for later nowcasts in these analyses, shorter pre-nowcast time series were used. For example, the 5-year nowcasts were based on 2003-2011 data, whereas the 1-year nowcasts were based on 2003-2015 data.

Figure 1.1 below shows the 3-year forward predictions (for 2016 alcohol-related hospital admissions) with corresponding 95% Bayesian Credible Intervals (95%bCI) for each of the 324 lower-tier local areas in the analyses.

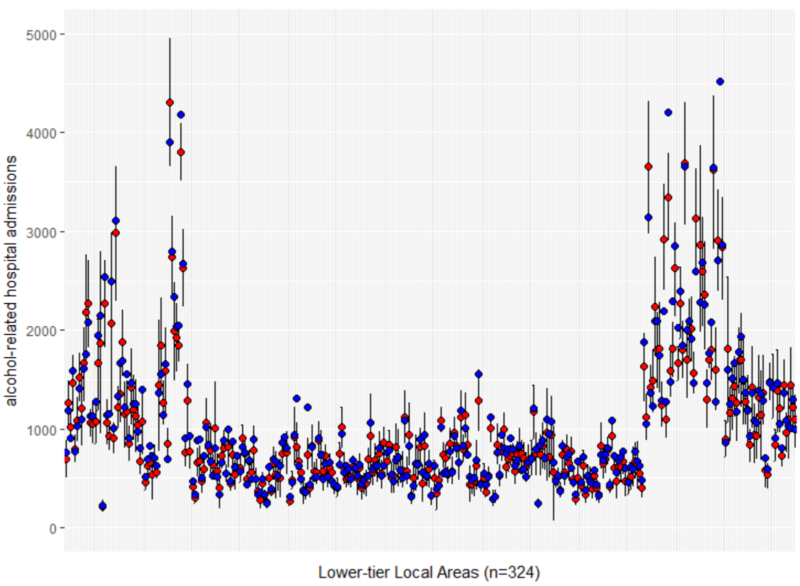


Figure 1.1 *Measured (blue) and predicted (red) alcohol-related hospital admissions for each lower-tier local area plus 95% Bayesian Credible Intervals, for 2016; predicted based on 2003-2013 time-series.*

The 3-year nowcasted hospital admissions show the same patterns as what was actually measured for 2016. Only a minority of these estimates were outside of the 95% bCI. These larger errors will be explored in detail below. Note however, that in this report we focus on accuracy of predictions, with the precision not being evaluated in detail.

A similar graphical overview for a 5-year nowcasted outcome (based on 2003-2011 timeseries) is shown in Figure 1.2; indicating comparable performance. In agreement with the results in Table 1, for 5-year nowcasts some individual estimates would have larger errors, and indeed several measured hospital admission counts (blue dots) can be observed outside of the 95% bCIs of the corresponding nowcasts (red dots).

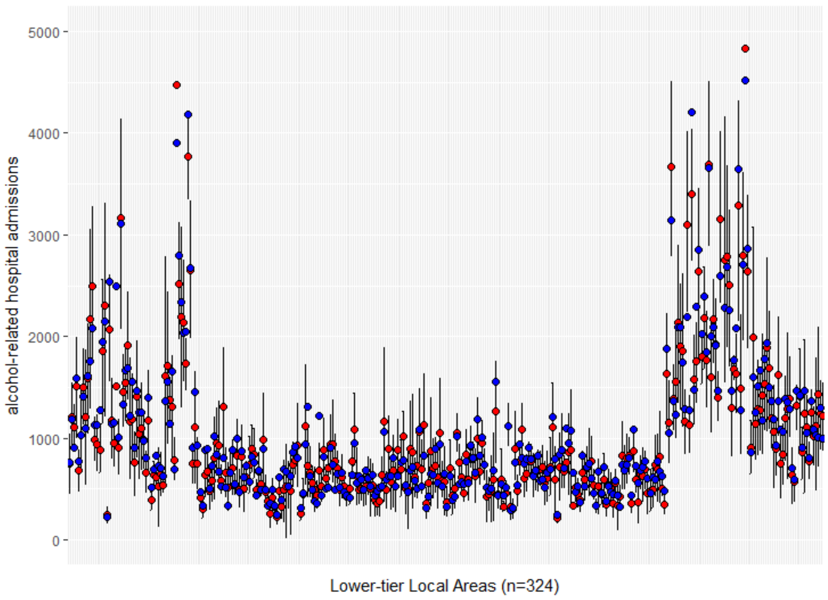


Figure 1.2 *Measured (blue) and predicted (red) alcohol-related hospital admissions for each lower-tier local area plus 95% Bayesian Credible Intervals, for 2016; predicted based on 2003-2011 time series.*

The distribution of individual 3- and 5-year prediction errors are shown in Figure 2 below. Errors ranged from -39.2% to +43.0% and -48.7% to 62.0% for 3- and 5-year nowcasting, respectively.

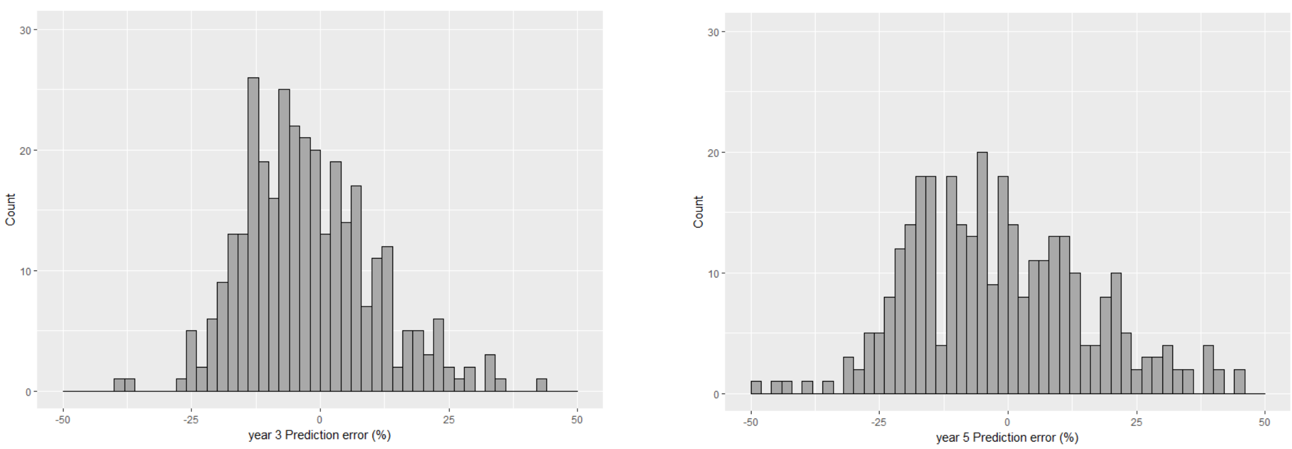


Figure 2. *Histogram of 324 relative prediction errors (in %) for 3 (left) and 5 (right) year predictions (2016, based on 2003-2013 and 2003-2011 time-series).*

Overall, for 3-year nowcasts, the models slightly underpredicted alcohol-related hospital admissions by 3.8% on average, with 50% of the errors relatively accurate between -11.7% (25th percentile) and +5.7% (75th percentile); i.e. the interquartile range (IQR). 174 of 324 estimates were within +/-10%, but only 9 within +/- 5%. In terms of absolute errors, these range from -859.6 to +735.4 admissions, with the IQR of absolute prediction errors from -94.5 admissions to +52.3 admissions.
Comparable performance was observed for 5 year predictions, with an average underprediction of 3.4%, and the IQR from -15.1% to +9.9%.

To assess what caused the largest prediction errors, the 6 areas (just to give an indication) with the highest absolute 3-year prediction errors were extracted, and the measured (straight line) and predicted (dotted) plus 95% bCI (l blue shading) time series are shown for graphical evaluation below in Figures 3.1 to 3.6.

Figures 3.1 to 3.6 *show measured (solid line), predicted (dashed line) and 95% Bayesian Credible Intervals (blue shading) alcohol-related hospital admission time series based on 2003-2013 (i.e. 1-10) modelling and 2014-2016 (i.e. 11-13) nowcasting.*

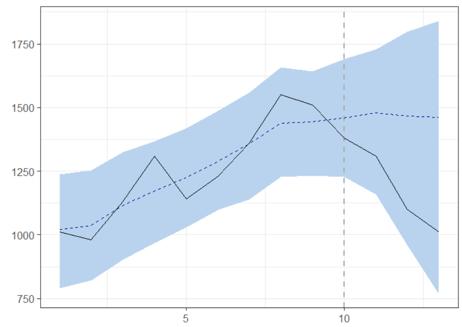


Figure 3.1. E09000031 (+43%)

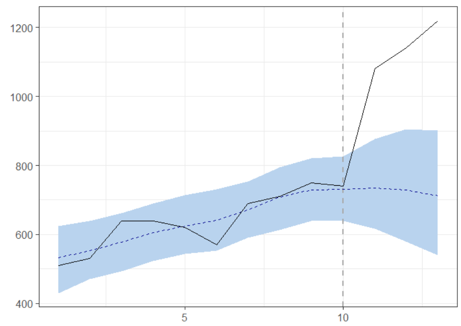


Figure 3.2. E07000076 (-39%)

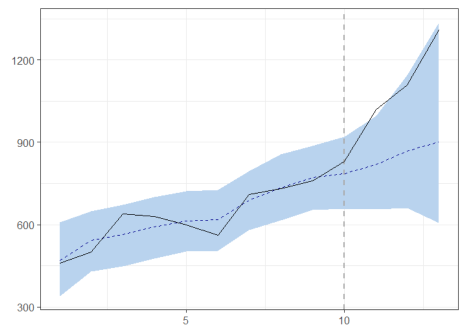


Figure 3.3. E07000071 (-37%)

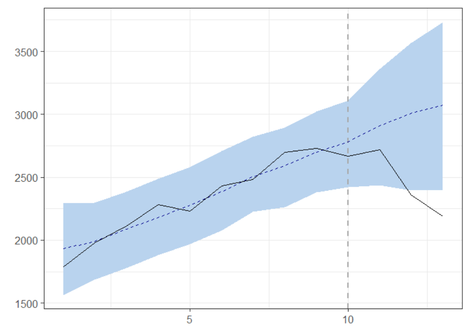


Figure 3.4. E07000112 (+36%)

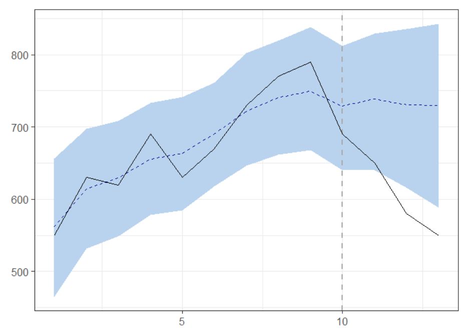


Figure 3.5. E07000138 (+34%)

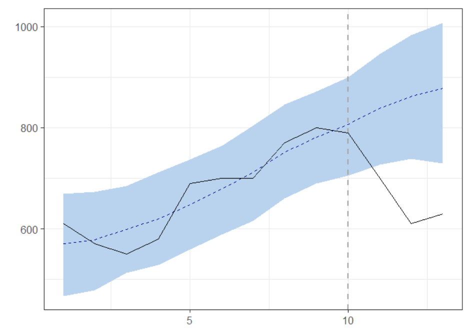


Figure 3.6. E08000010 (+34%)

As illustrated by Figures 3.1 to 3.6, the majority of the largest nowcasting errors results from an unexpected change in 2013, which, as indicated by the nowcasting is not related to demographic changes. Whether this is the result of a theoretically predictable change, for example due to policy changes, or whether these are due to measurement errors/bias, for example as a result of coding practices, is unknown. However, in the case of 324 individual and relatively short time-series models it is not unexpected that some nowcasts will have larger prediction errors because of changes that occur at the exact moment that nowcasting starts.

Different areas with largest prediction errors were found for 5-year nowcasting, but the patterns are similar. Again, the 6 areas with largest absolute errors are shown in Figures 4.1-4.6 below:

Figures 4.1 to 4.6 show measured (solid line), predicted (dashed line) and 95% Bayesian Credible Intervals (blue) alcohol-related hospital admission time series based on 2003-2011 (i.e. 1-8) modelling and 2012-2016 (i.e. 9-13) nowcasting.

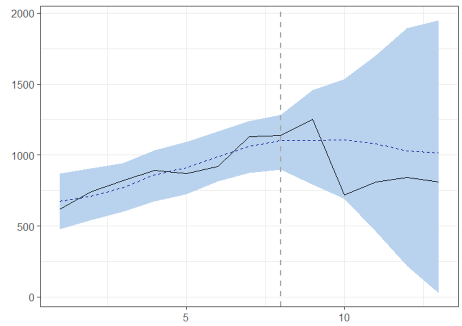


Figure 4.1. E07000028 (+62%)

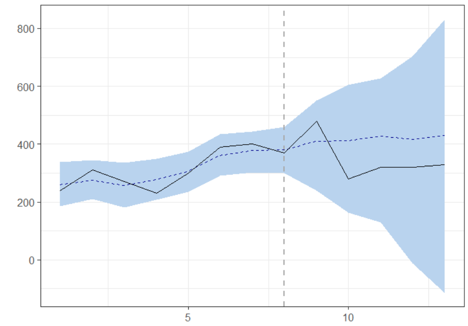


Figure 4.2. E07000030 (+57%)

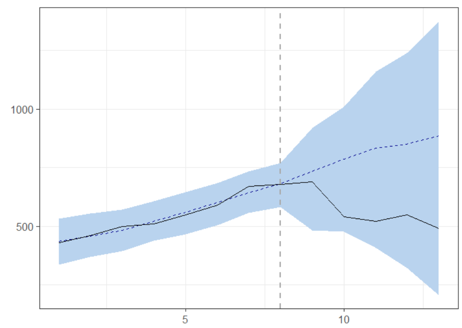


Figure 4.3. E07000083 (+57%)

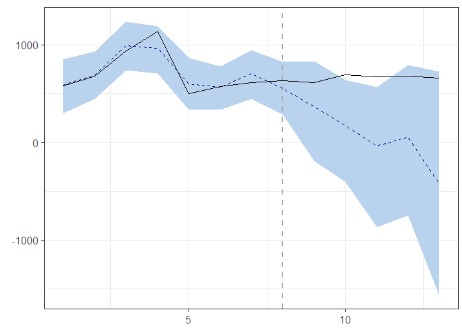


Figure 4.4. E07000198 (-49%)

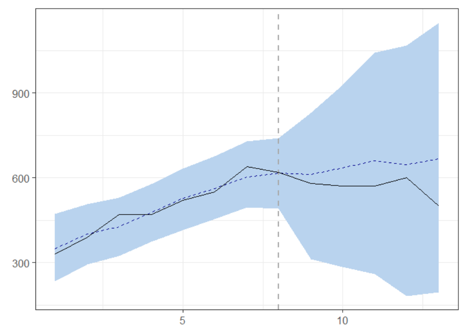


Figure 4.5. E07000120 (+46%)

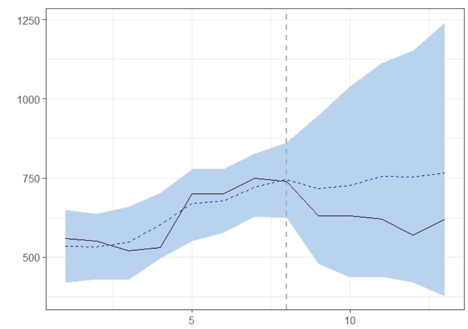


Figure 4.6. E07000108 (+45%)

Figure 5 shows the geographical distribution, at LTLA-level, of the absolute prediction errors for 5-year nowcasting; stratified in strata of 5% increase in errors. There is no clear geographical pattern of nowcasting errors.

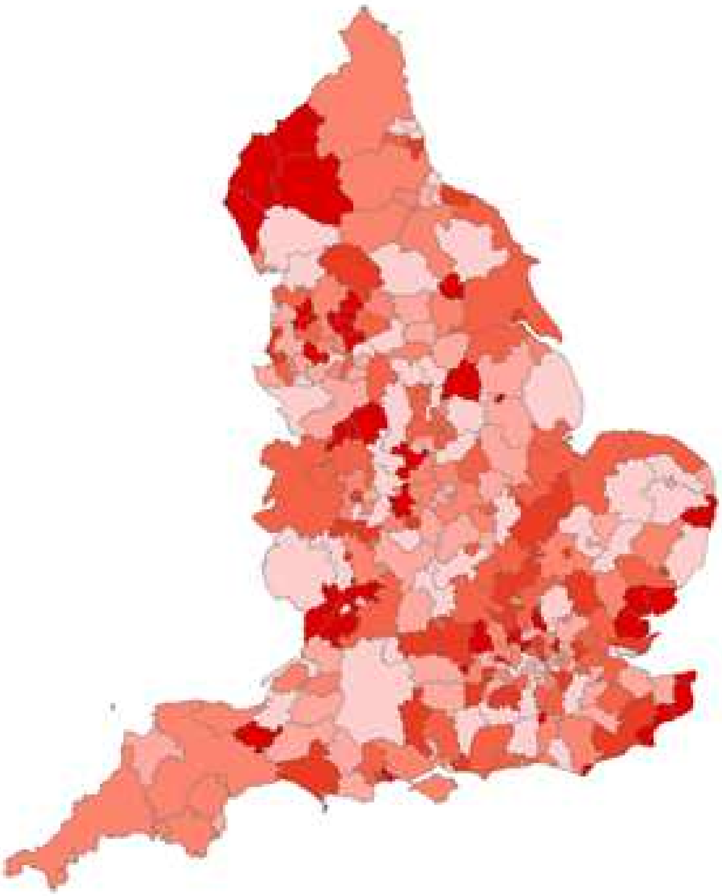


Figure 5. 5-yr prediction errors for each lower-tier local area (in percentage) stratified into: 0-5%, 5-10%, 10-15%, 15-20%, 20-25%, and 25%, with darker red indicating larger 5-year prediction error.

Despite the existence of prediction errors, 3- and 5-year nowcasting at the local level is overall of sufficient quality for meaningful prediction as shown in the Figures and Tables above, as well as in Figure 6 below. Figure 6 shows the overlap in the distributions of the 2016 5-year nowcasted values and the 2016 measured alcohol-related hospital admission counts, which are largely comparable. Least-squares linear regression of 5-yr (2016) nowcasted and measured LTLA alcohol related hospital admissions results in a regression coefficient of 0.98 and explained variance (R^2) of 98%, indicating very good fit.

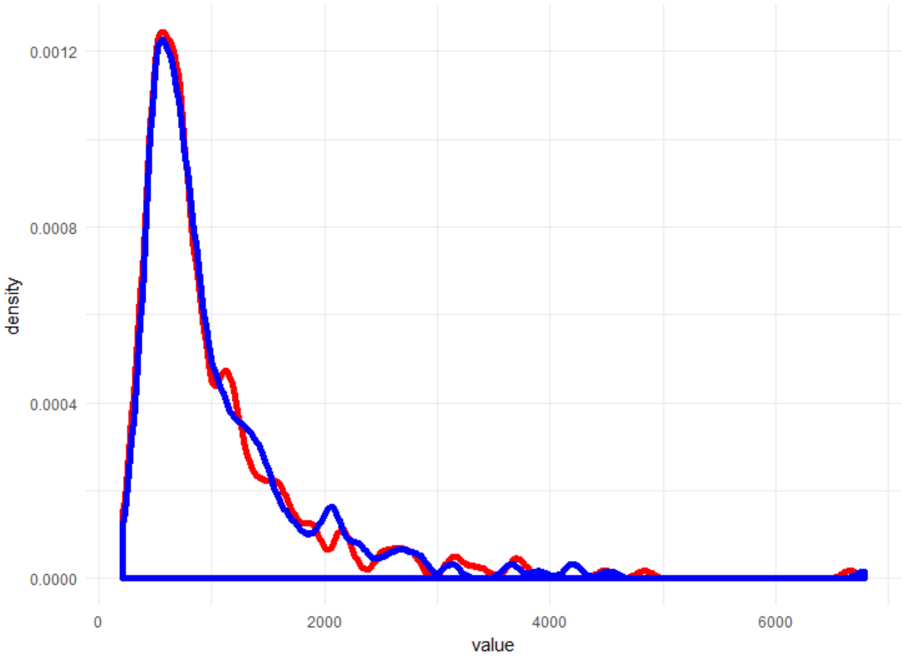


Figure 6. *Distribution, shown as a density function of measured (blue) and 5-year nowcasted (red) lower tier local area alcohol-related hospital admissions.*

A direct comparison of measured and 5-yr (2016) nowcasted alcohol-related hospital admissions are shown graphically in Figures 8.1 and 8.2.

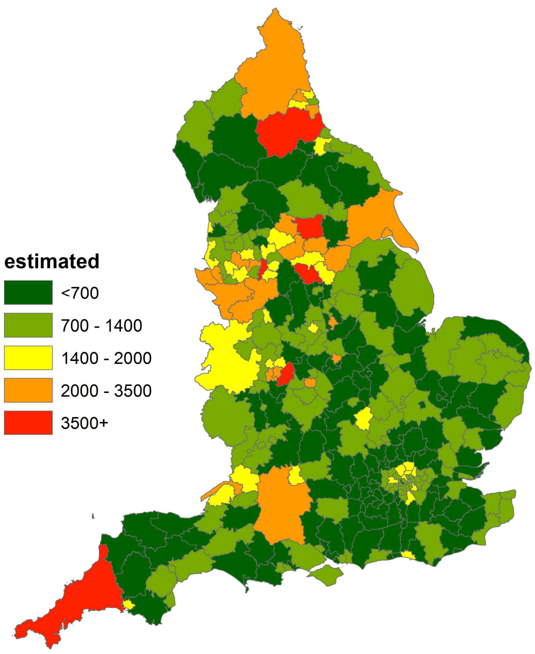


Figure 8.1. 5-year nowcasted, for 2016) alcohol-related hospital admissions for 324 lower-tier local authorities.

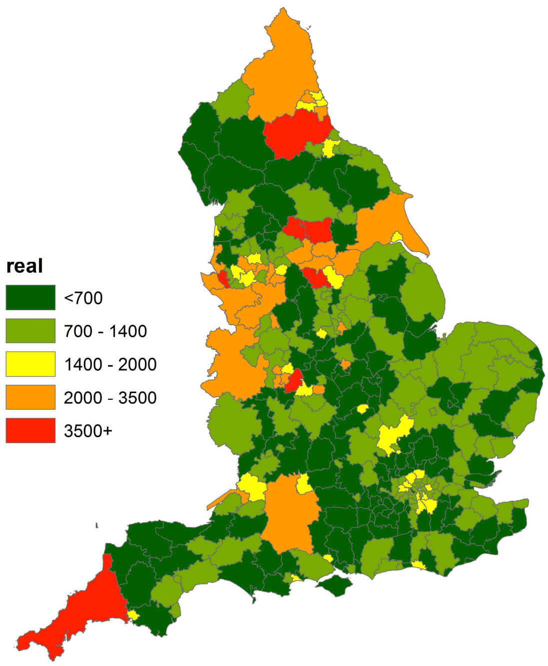


Figure 8.2. Observed 2016 alcohol-related hospital admissions for 324 lower-tier local authorities.

5-year forecasts (to 2021)

The same 324 models evaluated in the previous section were subsequently used to forecast annual alcohol-related hospital admission counts for 5 years to 2021. Forward projections were here based on the complete 2003-2016 timeseries, and used ONS-forecasted trends in each of the age groups for each LTLA.

The individual model results for each LTLA are shown graphically in the Appendix. In summary, the graphs indicate that the 2002/3 to 2015/16 temporal patterns are well modelled in the vast majority of the 324 models, as would be expected based on the evaluation of accuracy in the previous section. As can also be observed, and similar to the nowcasting assessment, there is variability in the accuracy of the modelling between the different LTLAs. There is some evidence that the errors are correlated with the variability in annual reporting of alcohol-related hospital admissions, as shown in Figure 9 (for 2016, so nowcasted rather than forecasted):

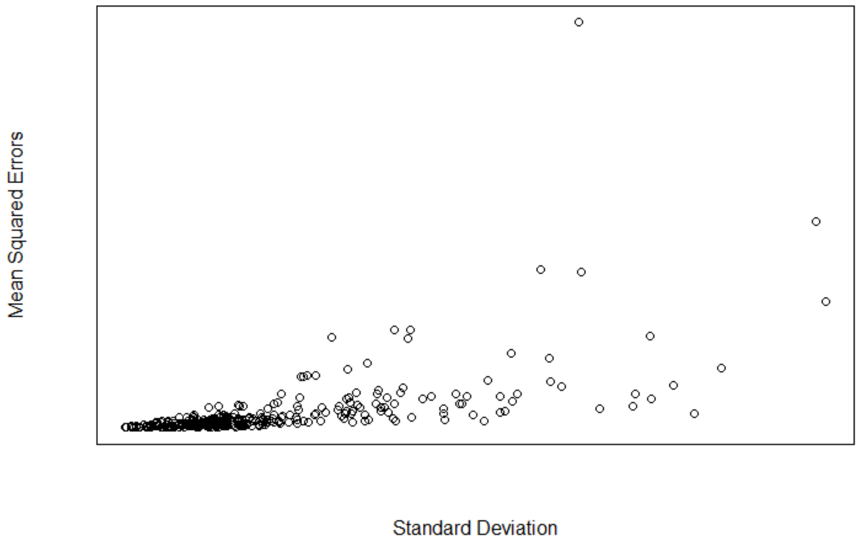


Figure 9. *Mean-squared errors of 2003-2016 measured and forecasted alcohol-related hospital admissions plotted against standard deviation of measured 2003-2016 alcohol-related hospital admissions. Linear least-squared means regression indicates a moderate positive association ($p<0.001$, adjusted R^2~36%).*

An overview of the 5-year forecasting model results are shown for different geographical aggregations in Table 2. At the national level, initial assessment of the accuracy of the modelling for 2002/3-2015/16 data indicates that the temporal trends are very well captured, with an average 2016 error of only -1.8%. Furthermore, the 5-year forecasts suggest that the total number of alcohol-related hospital admissions in 2021 is expected to increase slightly, by 1.3%, compared to 2016. However, given that the population increase is projected to be 3.7% over that same time period, the annual rate of alcohol-related hospital admission is projected to decrease (assessed in detail later in this report).

Table 2. Summary overview of 5-year (2021) forecasted alcohol-related hospital admissions.

	Total alcohol-related hospital admissions
England in 2016	
Measured (N=326)	
Measured (N=324)	336,330
Forecasted	330,136
Error (%)	-1.8%
England in 2021 (forecasted)	334,401
Forecasted change relative to 2016	+1.3%
Population England 2016 (in 1000s)	55,219
Projected in 2021 (in 1000s)	57,248
Change relative to 2016	+3.7%
Region	**Forecasted average* change**
East	-0.0%
East Midland	+4.6%
London	-9.1%
North East	-2.0%
North West	+5.7%
South East	+1.9%
South West	+2.7%
West Midlands	-0.3%
Yorkshire and The Humber	+0.6%
ONS Supergroup	**Forecasted average* change**
Affluent England	+0.6%
Business, Education and Heritage Centres	-4.0%
Countryside Living	+2.6%
Ethnically Diverse Metropolitan Living	-11.6%
London Cosmopolitan	-9.2%
Services and Industrial Legacy	+2.3%
Town and Country Living	+4.0%
Urban Settlements	+4.3%

*: *average of LTLA changes; not weighted by population size of LTLA*

LTLA-specific estimates of the 5-year change are shown in Figure 10:

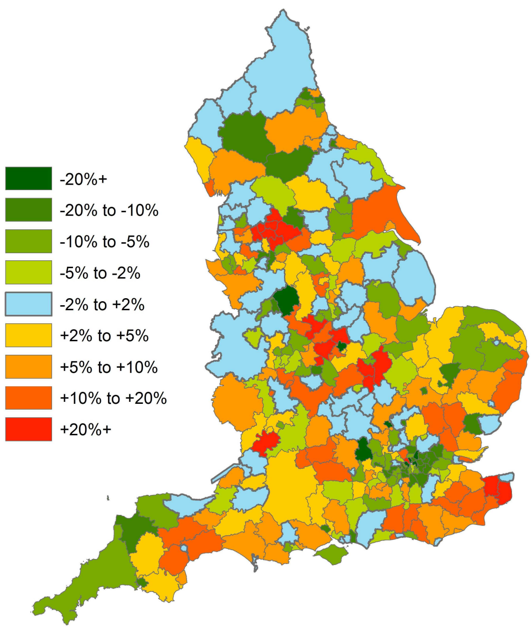


Figure 10. *5-yr forecasted (to 2021) change from the alcohol-related hospital admission rate compared to the final measured year (2016) for all LTLAs. Exact numbers are provided in the Appendix.*

We evaluated whether there is evidence of regression-to-the-mean in our forecasted predictions by plotting the last measured alcohol-related hospital admission rate (in 2016) against the percentage change in 2021. Least-squares regression indicates there is some evidence of regression-to-the-mean (β~-0.003; $p<0.001$; adjusted R^2~4%), and the association is shown graphically in Figure 11:

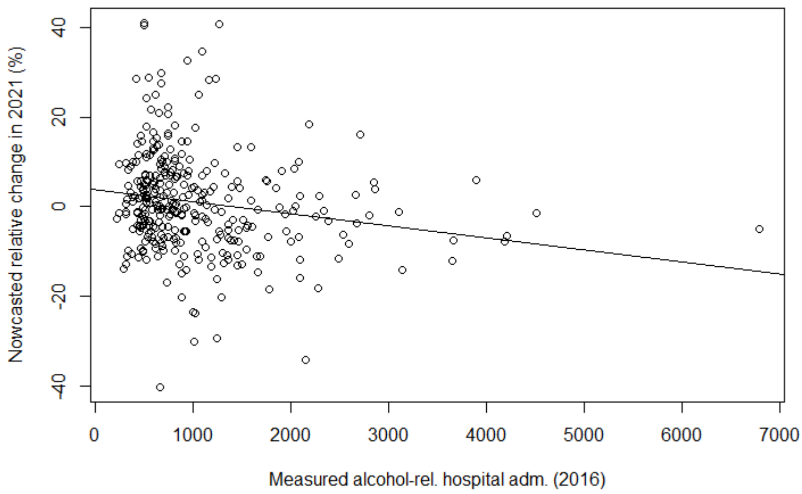


Figure 11. *Correlation between measured lower-tier local authority alcohol-related hospital admissions in 2016 and forecasted relative change in 2021.*

Aggregation at the regional level shows relatively large differences between the regions of England, with the largest increases in alcohol-related hospital admissions forecasted to be in the North West and East Midlands and the lowest in London, respectively (Table 2 and Figure 12).

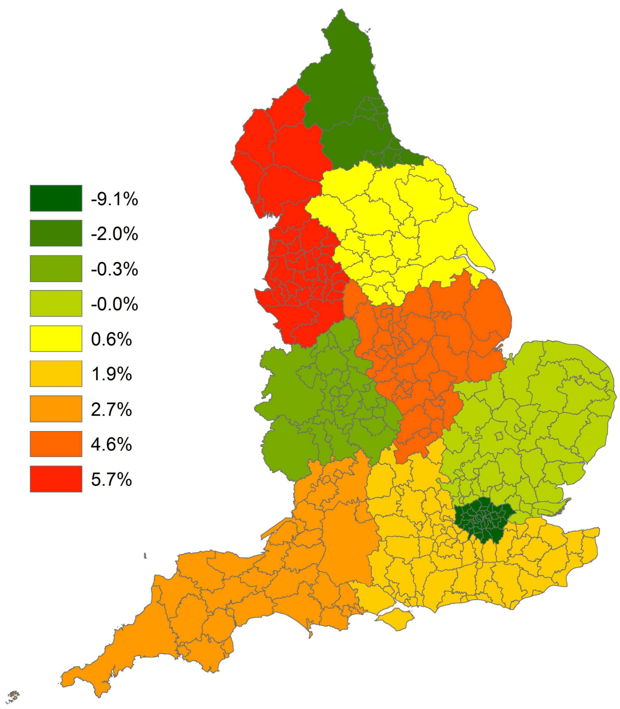


Figure 12. *5-yr forecasted (to 2021) change from the alcohol-related hospital admission rate in the final measured year (2016) averaged by ONS Regions. Refer to Table 2 for Regions.*

Aggregation by ONS Supergroup (Table 2 and Figure 13) shows a more nuanced picture, forecasting that the largest decreases are expected to occur in the Ethnically Diverse Metropolitan and London Cosmopolitan areas, while the largest increases in hospital admissions are forecasted in areas of Town and Country Living and Urban Settlements.

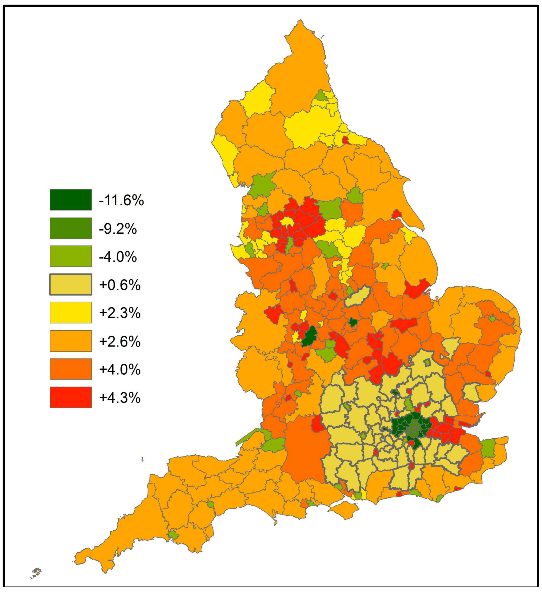


Figure 13. *5-yr forecasted (to 2021) change from the alcohol-related hospital admission rate in the final measured year (2016) averaged by ONS Regions. Refer to Table 2 for ONS Supergroups.*

The ONS Grouping is a hierarchical classification, and can be further deconstructed to 'Groups' and 'Subgroups'. One level down from the Supergroup, aggregation at Group-level, is shown in Figure 14, and shows that the largest reduction in hospital admissions are expected in the ethnically diverse metropolitan and London cosmopolitan areas, but also in University towns and cities. The largest forecasted increases however, are observed in remote costal living and country living areas, but especially in the areas with manufacturing traits.

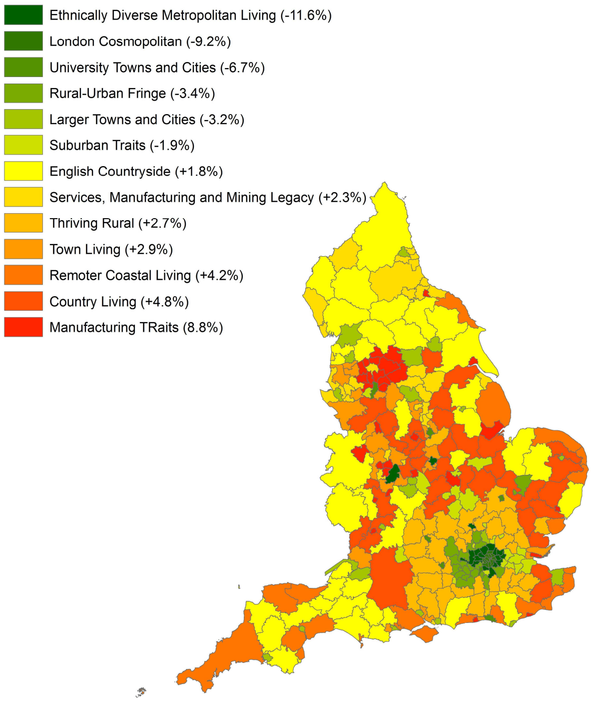


Figure 14. *5-yr forecasted (to 2021) change from the alcohol-related hospital admission rate in the final measured year (2016) averaged by ONS Group.*

Further stratification (and averaging) to the lowest Group-level (ONS Subgroup) is shown graphically in Figure 15.

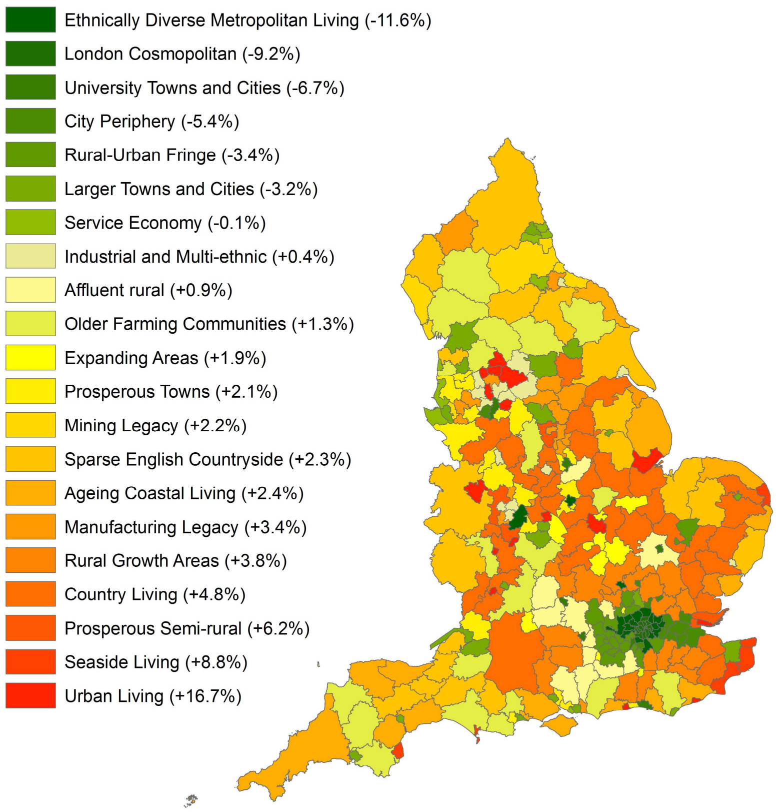


Figure 15. *5-yr forecasted (to 2021) change from the alcohol-related hospital admission rate in the final measured year (2016) averaged by ONS Subgroup.*

The largest reductions in alcohol-related hospital admissions are to be expected in ethnically diverse metropolitan living areas (-11.6%), London cosmopolitan (-9.2%) and University towns and cities (-6.7%), while the largest increases are expected in semi-rural (+6.2%), seaside (+8.8%), and urban living (+16.7%) areas.
We can hypothesize about these expected trends, in that illustrate two trends, which have been shown for alcohol-related mortality in England and Wales previously (Erskine et al., 2010):

1) increase in alcohol-related hospital admissions in the more deprived inner-city areas across the country, with the notable exception of London, and decreases in areas with with Universities and a large service economy.
2) decreases in geographical areas associated with movement of younger people, such as Metropolitan London, larger University towns and cities and city peripheries; and increases in geographical areas assocated with the movement of older people such as to rural and coastal areas at middle or higher age.

Both of these trends are consistent with literature, and describe how:

1) alcohol harms are disproportionally experienced by lower socio-economic groups, and
2) how harms are not necessarily experienced where consumption is highest, with these predictions indicating alcohol-related health harms experienced at later age, and thus disproportionally in the aging populations in (semi-)rural and coastal communities.

Figure 16 shows the results of further exploration of the first hypothesis:

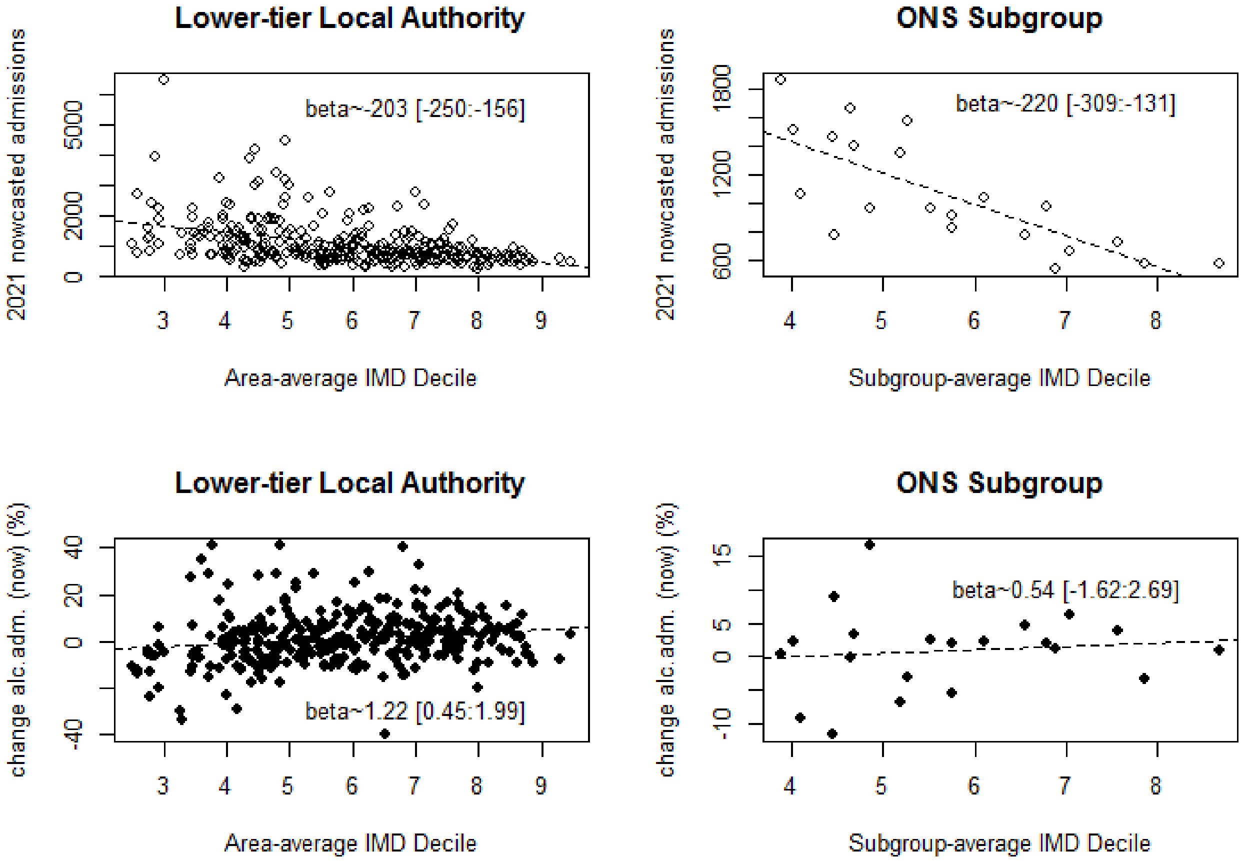


Figure 16. *Correlations between Index of Multiple Deprivation (IMD; lower indicates more deprived) decile and forecasted 2021 hospital admissions (top row) and 2016-2021 change (bottom row).*

As is already well-known, there is a clear association between social deprivation, measured as the Index of Multiple Deprivation (IMD) score, and alcohol-related hospital admission rates, and this is expected to remain the case in 2021 (top two figures in figure 16). However, the forecasted change from 2016 to 2021 does not show a similar association, and if anything, increase in admissions will be slightly higher (compared to 2016) in less deprived areas (bottom two figures in Figure 16).

To explore the 2nd hypothesis further, Figure 17 shows the average forecasted change in alcohol-related hospital admissions from 2016 to 2021 (in %) plotted against the average forecasted population change (in %) over that same period, for each of the 21 ONS subgroups.

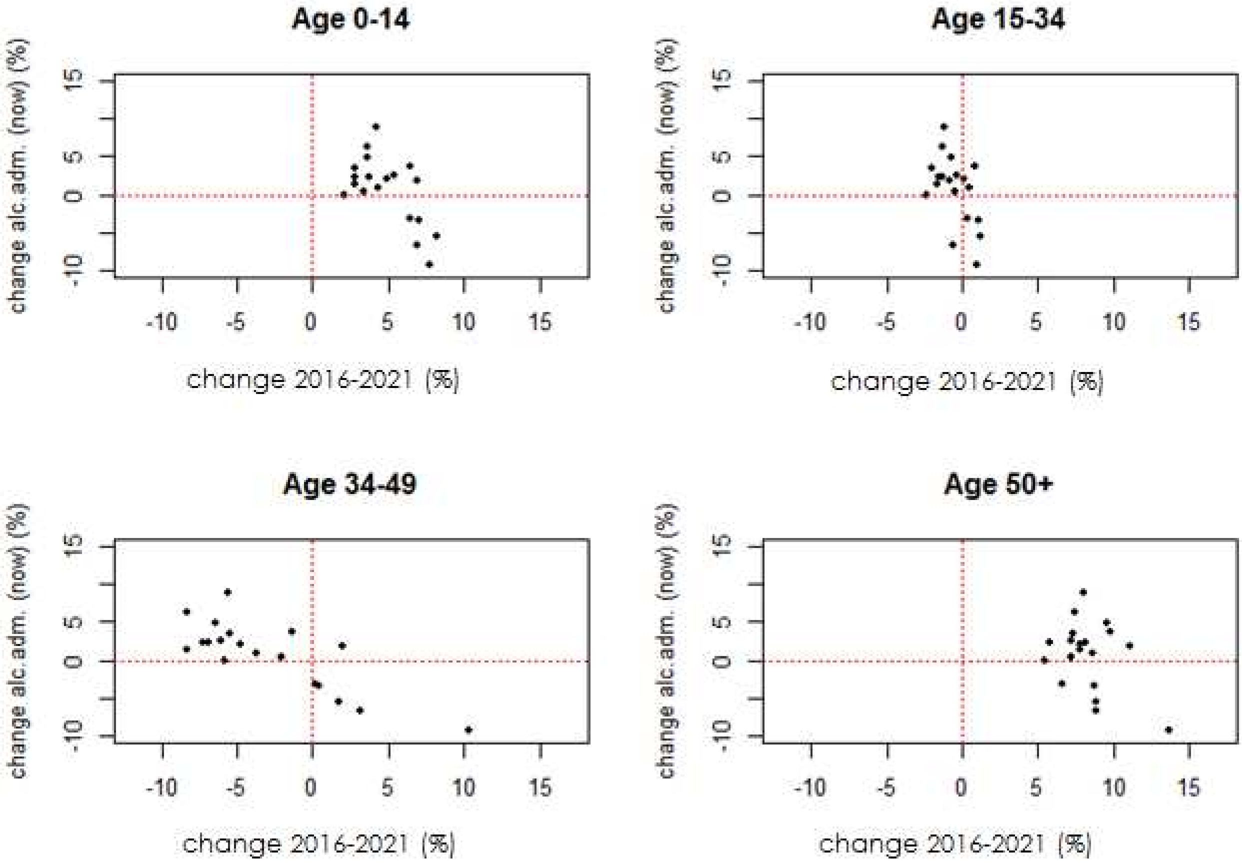
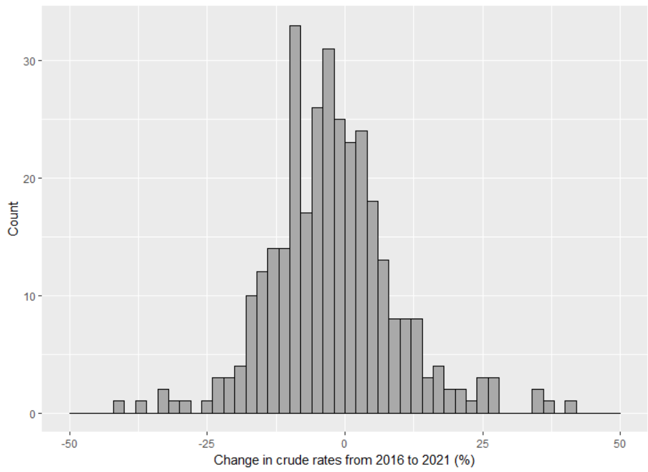


Figure 17. *correlation between change in 2016-2021 population size and alcohol-related hospital admissions.*

As shown, indeed, in areas with with a forecasted increase in the 50+ age groups (and 0-14, but that is less relevant), but a decrease in the age group 34-49 and to a lesser extent 15-34, alcohol-related hospital admissions are forecasted to increase in 2021 (compared to 2016).

There are, of course, other hypothesizes that can be explored in relation to these forecasted trends, such as for example whether trends in the make-up of ethnic groups in areas may also contribute to the observed variation, but further exploration was not part of the current 'proof of principle' study, and we plan to investigate this at a later stage.

In addition to the local number of alcohol-related hospital admissions and the forecasted change therein (which is important to anticipate local healthcare costs for example), the forecasted burden, measured as rates (per 100,000 population), is also important. Therefore, crude rates of alcohol-related hospital admissions were calculated for each (n=324) LTLA for 2016 and 2021.
The first important result of these analyses is that nationally, the crude incidence rate of alcohol-related hospital admissions is expected to reduce by 2.3% in 2021. This does, however, range from -40.8% to +41.6%, with 50% of LTLAs between -9.2% (25th percentile) and +3.8% (75th percentile). The distribution in forecasted 2016-21 rate changes is shown below in Figure 18 and its geographical variation in Figure 19:



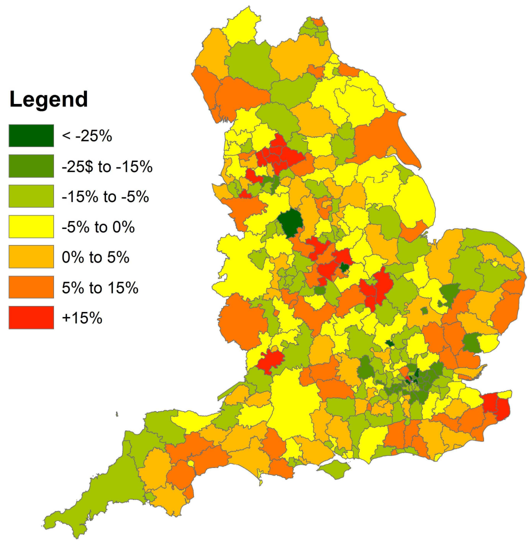
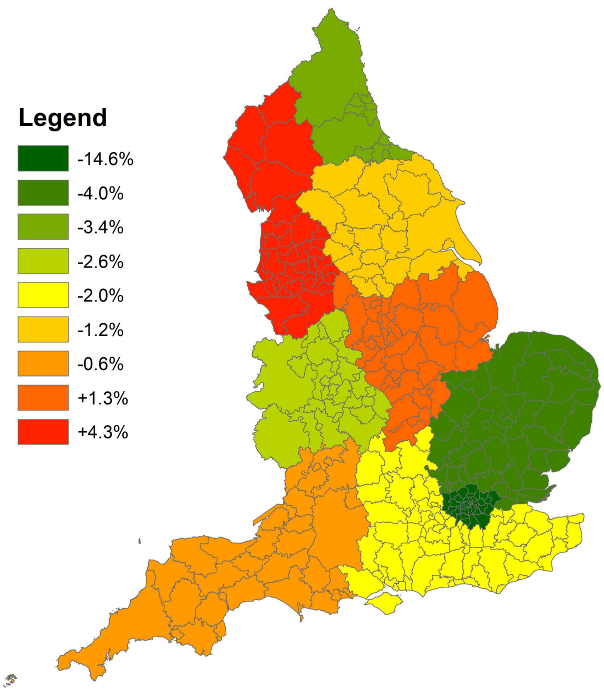


Figure 18. *Histogram of the forecasted 2016-21 changes in crude alcohol-related hospital admission rates for 324 lower-tier local authorities.*

Figure 19. Forecasted changes in the crude *alcohol-related hospital admission rates from 2016 to 2021 for 324 lower-tier local authorities.*

Similar to the analyses done for counts, geographical aggregation at different levels can be done for rates as well (Figures 20-22):



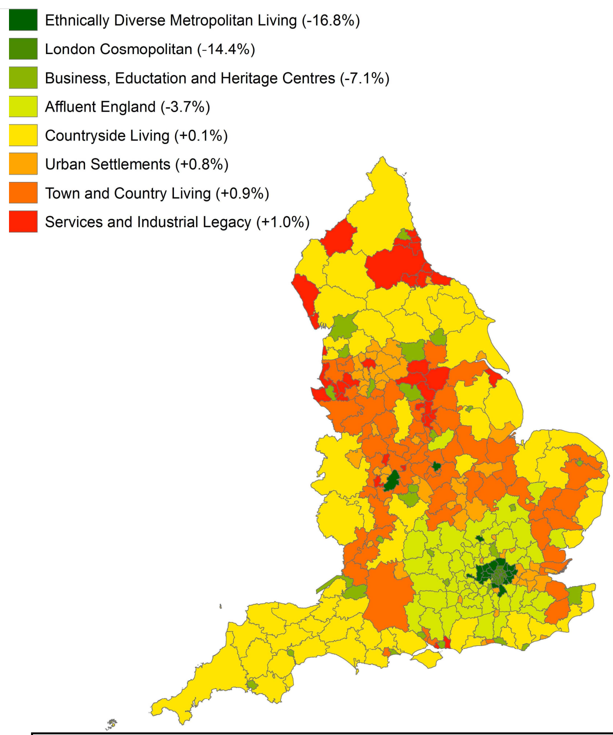


Figure 20. *Forecasted 2016-21 changes in crude alcohol-related hospital admission rates aggregated by Region.*

Figure 21. *Forecasted 2016-21 changes in crude alcohol-related hospital admission rates aggregated by ONS Group.*

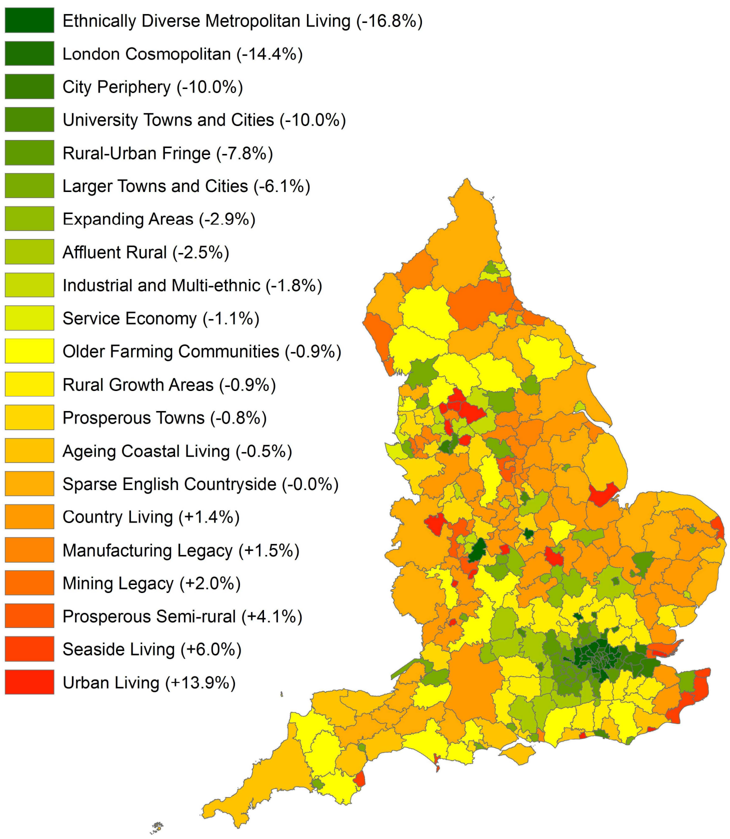


Figure 22. *Forecasted 2016-21 changes in crude alcohol-related hospital admission rates aggregated by ONS Subgroup.*

The results are comparable to those forecasted for the absolute counts as a result of relatively small 5-year changes in LTLA population sizes. However, the numbers have changed slightly, with the forecasted reduction in annual alcohol-related hospital admission rates in 2021 (relative to 2016) being somewhat larger (for example, -17% forecasted in Ethnically Diverse Metropolitan Areas) and in contrast, the maximum increase somewhat lower (i.e. urban living +14%).

Example: impact of alcohol taxation

As an example of how the above framework can be used to now- or forecast specific scenarios, we here use it to assess the impact of alcohol taxes. This work complements previous studies on the (potential) impact of specific taxation policies (For example: Sheron and Gilmore, 2016; Meier et al., 2016). Forecasts are based on the 2003-2021 time-series of real alcohol duty rates (in real terms, relative to 1978) for beer, wine and spirits obtained from The IFS Green Budget 2016 (Figure 9.10), shown in Figure 17.

Note that these are national rates and forecasts, rather than area-specific changes, and therefore the same 3 time-series have been added to the datasets for each of 324 Bayesian structural time series models.

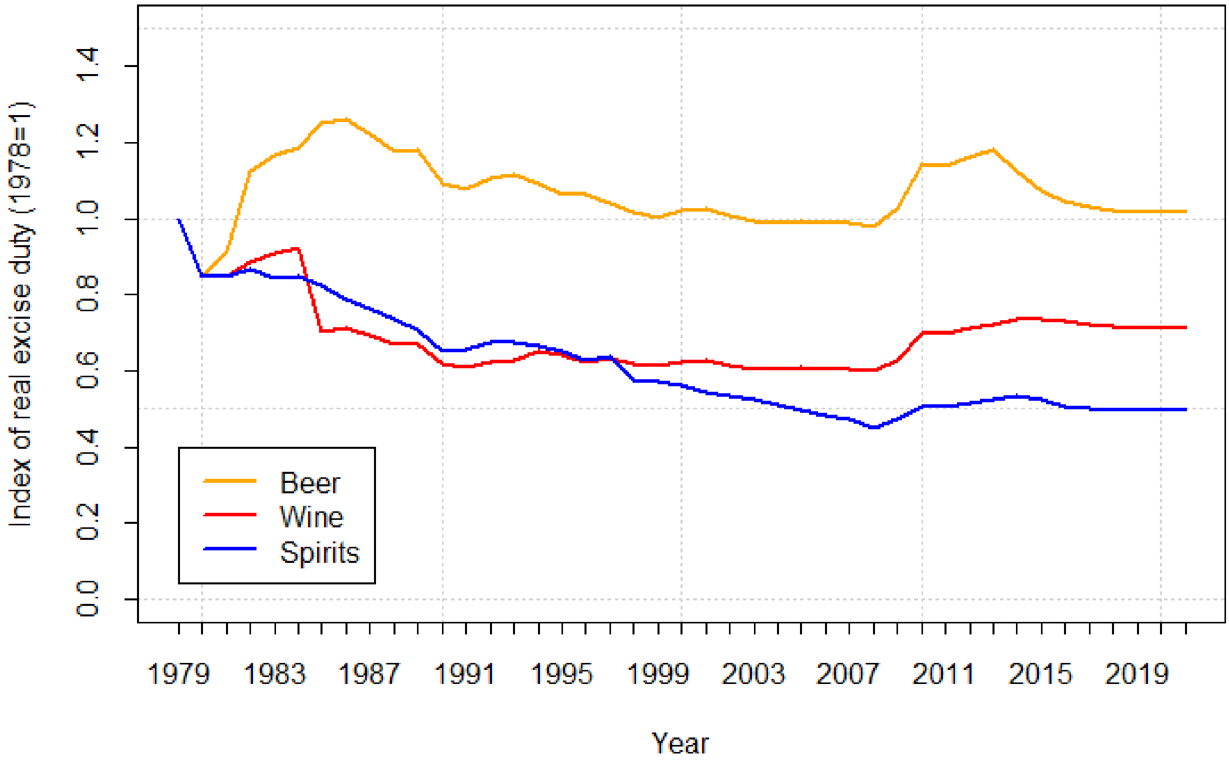


Figure 23. *Real alcohol duties, 1978-79 to 2020-21. Note: Assumes beer at 3.9% ABV, wine not exceeding 15% ABV and spirits at 40% ABV. Rates are indexed relative to levels in 1978–79. Dashed lines indicate announced future policy.*
Original source: HMRC website, https://www.gov.uk/government/organisations/hm-revenue-customs; HM Treasury, Tax Benefit Reference Manual 2002–03 Edition, 2002; various HMRC / HM Customs & Excise Annual Reports.
Original data obtained from the author: Peter Levell

The results are shown in Table 3. As shown, the inclusion of additional explanatory variables – in this case real excise duties of beer, wine and spirits – leads to small improvements in accuracy in the order of 0.5% for average error, on average. Mean-squared errors (MSEs) also improved, with reductions in MSE indicating improved accuracy of predictions with increased temporal distance.
Interestingly, the 3-year forecasts on average underpredict true alcohol-related hospital admissions a bit more (-4.7%), but the range of 3-year prediction errors is smaller (-41.8% to 37.2%).

Table 3. *Accuracy of forecasting of 2014-2016 alcohol-related hospital admissions (based on 2003-2013 measured data). Data particle replicated from Table 1.*

	Forecasting 1 year	2 years	3 years
National-level alcohol related hospital admissions			
Measured	326,940	326,950	336,330
Forecasted	325,971	327,080	327,952
Forecasted including tax	319,856	321,141	323,130
Accuracy	97.8%	98.2%	96.1%
LTLA-level alcohol related hospital admissions			
Average	1,009	1,009	1,038.
Absolute average error	64.3	82.7	107.9
% error	6.4%	8.2%	10.4%
MSE	8693	15,959	25,589
Forecasting including tax			
Average	1,009	1,009	1,038
Absolute average error	63.7	78.9	101.1
% error	6.3%	7.8%	9.7%
MSE	8,075	13,241	20,918
MSE improvement	+7%	+17%	+18%

The difference of the 3-year forecasted alcohol-related hospital admissions with and without the inclusion of 'real alcohol excise duties' can be interpreted as the impact of changes in these taxes on alcohol-related hospital admissions. This is summarised in Table 4:

Table 4. *Relative difference in alcohol related hospital admissions as a result of forecasted changes in 'real alcohol excise duties.*

	Relative difference in alcohol related hospital admissions*
National mean	-0.2%
Maximum reduction at area level	-19.0%
25% percentile	-4.2%
75% percentile	+3.0%
Maximum increase at area level	+30.1%

**: relative to predictions from models not including tax variables, and unweighted with respect to area population size.*

The results indicate that the changes in taxation, which over the forecasting period roughly correspond to a reduction in real terms in alcohol tax for beer accompanied by an increase in those for wines and spirits, may be associated with a national public health effect of a reduction in alcohol-related hospital admissions of 0.2%, or in absolutely numbers has averted 2,713 admissions nationally in 2016. However, this differs widely between local areas, and in some areas, alcohol taxation was associated with unexpected increases (compared to the forecasted trend without taxation in the forecasting models).

A series of spatial maps is shown below, showing how the relative difference in alcohol-related hospital admissions, related to changes in real tax, varies across England (Figure 24-26):

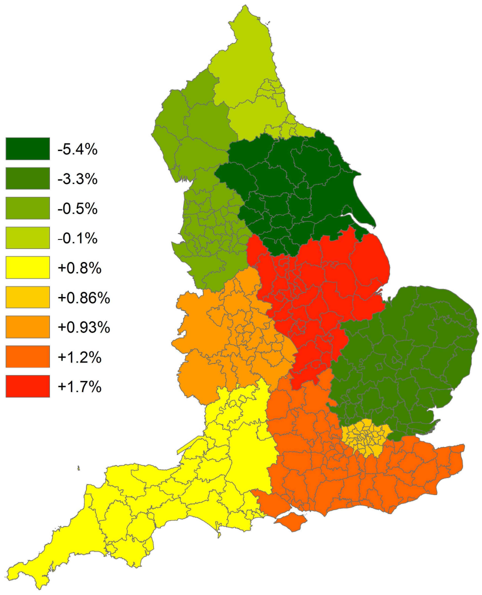


Figure 24. *Forecasted 2021 effect on alcohol-related hospital admissions associated with alcohol taxation forecasted trend. Aggregated by Region.*

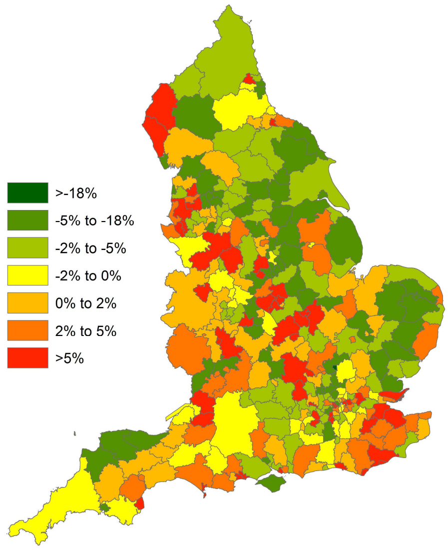
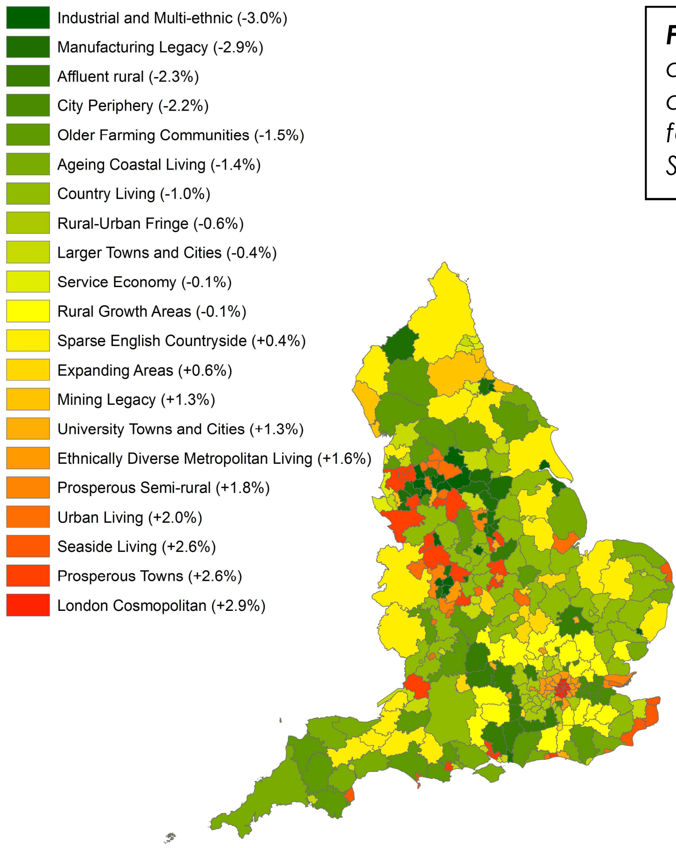


Figure 25. *Forecasted 2021 effect on alcohol-related hospital admissions associated with alcohol taxation forecasted trend. For each LTLA*

Figure 26. *Forecasted 2021 effect on alcohol-related hospital admissions associated with alcohol taxation forecasted trend. Aggregated by ONS Subgroup.*



Interestingly, as shown in Figures 24-26, the effect of forecasted real-term alcohol taxation trends for 2016-2021 are forecasted to have a larger positive effect (as in, resulting in a reduction in alcohol-related hospital admissions) in the North than in the South and, importantly, have the highest forecasted reductions in the industrial and multi-ethnic areas and areas with a manufacturing legacy. The highest forecasted increases instead, are found in prosperous towns and London cosmopolitan areas.

As such, these forecasts suggest that taxation is likely to have a larger impact in more deprived areas than in less deprived areas, which confirms a priori expectation.

CONCLUSIONS

This exploratory project has shown that using local area-level time series can be used to model time series of alcohol-related hospital admissions based on trends in age groups alone with reasonably good accuracy and precision, and as a result will allow for the forecasting of future trends in alcohol-related admissions. Forecasts can also explore impact of other factors as long as data on trends are available. Broadly, now- and forecasted trends in hospital admissions across England are in agreement with hypothesized trends, including the conclusion that alcohol taxation (in real terms) has a larger effect on deprived than on non-deprived communities.

Main Conclusions

As such, the main conclusions from these specific analyses are:

- The Bayesian structured time-series approach provides relatively accurate estimates of time trends of alcohol-related hospital admissions based on demographic changes alone, with an average 5-year forecasting error of about -2% (absolute error ~13%).
- At the national level, the annual number of alcohol-related hospital admissions is forecasted to increase by an extra 4,265 by 2021. However, because the population size is forecasted to increase more, this in relative terms corresponds to a reduction in the crude rate (per 1,000 people) of 2.3%.
- Increases in hospital admissions differ between regions, local areas, and types of local areas, and are expected to increase most (relative to 2016) in the North West and least in the London Cosmopolitan areas.
- Aggregation by type of LTLA suggests the largest increases are forecasted in the deprived but also the prosperous semi-rural and seaside living areas. The largest decreases are forecasted in metropolitan and ethnically diverse cosmopolitan areas and University towns and cities. These trends seem related to harms from alcohol disproportionally affecting more deprived communities and a secular trend of the movement of middle-aged and older people away from cities to (semi-) rural and seaside towns.

The 5-year forecasted impact from forecasted real alcohol taxation on alcohol-related hospital admissions is, on average, small but positive (in terms of hospital admissions), with a -0.2% decrease in expected admissions. Alcohol taxation is expected to lead to a greater reduction in alcohol-related hospital admissions in areas with higher levels of deprivation.

Limitations and Strengths of the methodology

This modelling approach has some limitations to keep in mind when utilizing it for forecasting purposes:

i. Although the local-area specific modelling we did in this project has the distinct benefit of creating individual models with confounding effects included, in terms of understanding causal associations between hospital admissions, demographic changes and confounding factors, these are largely treated as a 'black box'.
ii. As with any forecasting method, the quality of the forecasts is highly dependent on the quality of the data. This has not been evaluated in this study, but was assumed to be without measurement error.
iii. Although the accuracy of the estimates, which we deemed most important with respect to forecasting, is relatively good, we have not evaluated the precision of the forecasting in great detail here.

However, compared to other statistical methods there are several of distinct strengths to the use Bayesian structural time -series:

i. In contrast to normal, static, regression generally used, this methodology uses weighted inclusion of other model factors based on their inclusion probabilities and evolving model parameter estimates, and as such better represent the effects of covariates over time;
ii. The additional benefit of the weighting of covariates is that many factors can be included to estimate the time series without stronger prior hypotheses, since non-relevant parameters will be down-weighted or excluded
iii. The model does not rely on assumptions of linearity in that, if the trends in other factors is not linear this will be taken into account (and thus it does not rely on estimation of one linear "average" trend")

iv. the Bayesian framework allows for the incorporation of prior knowledge, which will enable more accurate representation of the true certainty of the estimates, and will evolve (improve its accuracy) over time.

And additionally, for the methodology used in this project, additional strengths are:

i. The analyses at lower-tier local area level allows the specific evaluation at local level for, say policy making, but in addition enables aggregation at any geographical level to draw more general inferences that could be helpful in planning policies. For example, in this report we observed that the largest increases in alcohol-related hospital admissions are likely in poorer urban areas, but also in prosperous semi-rural and seaside living areas, solely as a result of demographic changes.

ii. The associations between demographic changes and alcohol-related hospital admissions will be potentially confounded by a large number of other factors, such as deprivation, which will differ between different communities. Although, in theory, these could be specifically modelled if the data were available, the local-area-specific modelling we used in this study allows different associations (in terms of magnitude and inclusion of specific associations) to be modelled for each area, thereby taking into account the effects of confounding factors automatically and thus greatly improves the accuracy of these now- and forecasts. This has a drawback as well however in that this considers the causal model largely as a 'black box', which was mentioned above.

DISSEMINATION AND FUTURE WORK

The outputs of the project, other than this final report, are expected to be:

I. A project-specific 'Alcohol Insight' report will be written once the final report is approved.

II. We have started drafting the work for submission to a peer-reviewed scientific journal, and have identified the 'International Journal of Epidemiology' as the initial target journal.

III. An abstract has been submitted to the international Kettil Bruun Society for the Social and Epidemiological Research on Alcohol annual conference to facilitate faster dissemination to the international scientific community. We are looking into possibilities of presenting the work at a national conference this year to target local and national stakeholders.

IV. As per the research proposal, an example R script has been provided in the Appendix to this report, which enables interested stakeholder to explore the methodology without the requirement of extensive programming in the R syntax.

Future work will include assessment of the association between specifications of the models and precisions of the estimates and forecasts initially, while we also aim to make the framework more user-friendly. This methodology is likely to be also very useful for forecasting purposes of other outcomes, which may not be related to alcohol-harms. Possibly, and this will be further discussed, this work can be done in future research grant applications.

REFERENCES

Bańbura M GD, Reichlin, L. Nowcasting: ECB Working Paper, No. 1275. 2010: http://hdl.handle.net/10419/153709.

Brodersen KH, Gallusser F, Koehler J, Remy N, Scott SL. Inferring causal impact using Bayesian structural time-series models. The Annals of Applied Statistics. 2015;9(1):247-74.
Brodersen KH, Hauser A. Package 'CausalImpact'. 2017. https://google.github.io/CausalImpact/

Chen Z, Brown EN. State space model. Scholarpedia. 2013;8(3):30868. doi:10.4249/scholarpedia.30868

Cowles MK, Carlin BP. Markov chain Monte Carlo convergence diagnostics: a comparative review. Journal of the American Statistical Association. 1996;91(434):883-904.

de Almeida Marques-Toledo C, Degener CM, Vinhal L, Coelho G, Meira W, et al. (2017) Dengue prediction by the web: Tweets are a useful tool for estimating and forecasting Dengue at country and city level. PLOS Neglected Tropical Diseases 11(7): e0005729. https://doi.org/10.1371/journal.pntd.0005729

de Vocht F, Tilling K, Pliakas T, Angus C, Egan M, Brennan A, et al. The intervention effect of local alcohol licensing policies on hospital admission and crime: a natural experiment using a novel Bayesian synthetic time-series method. Journal of Epidemiology and Community Health. 2017a: jech-2017-208931.

De Vocht F, Heron J, Campbell R, Egan M, Mooney J, Angus C, et al. Testing the impact of local alcohol licencing policies on reported crime rates in England. Journal of Epidemiology and Community Health. 2017b;71(2):137-45.

Erskine S, Maheswaran R, Pearson T, Gleeson D. Socioeconomic deprivation, urban-rural location and alcohol-related mortality in England and Wales. BMC Public Health. 2010;10(1):99.

Giannone D, Reichlin L, Small D. Nowcasting: The real-time informational content of macroeconomic data. Journal of Monetary Economics. 2008;55(4):665-76.

Lampos V, Miller AC, Crossan S, Stefansen C. Advances in nowcasting influenza-like illness rates using search query logs. Scientific reports. 2015;5:12760.

Meier PS, Holmes J, Angus C, Ally AK, Meng Y, Brennan A. Estimated effects of different alcohol taxation and price policies on health inequalities: A mathematical modelling study. PLOS Medicine. 2016;13(2):e1001963.

Meng X-L. Posterior predictive p-values. The Annals of Statistics. 1994:1142-60.

Met Office 2011. Nowcasting. https://www.metoffice.gov.uk/learning/making-a-forecast/hours-ahead/nowcasting.

Office for National Statistics 2012. https://www.ons.gov.uk/ons/guide-method/method-quality/specific/public-sector-methodology/articles/nowcasting-healthcare-output.pdf

Office for National Statistics, 2016. https://www.ons.gov.uk/economy/economicoutputand productivity/publicservicesproductivity/methodologies/newnowcastingmethodsformoretimely quarterlyestimatesofuktotalpublicserviceproductivity

Sheron N, Gilmore I. Effect of policy, economics, and the changing alcohol marketplace on alcohol related deaths in England and Wales. BMJ. 2016;353.

Scott SL, Varian HR. Predicting the present with bayesian structural time series. International Journal of Mathematical Modelling and Numerical Optimisation. 2014;4-23.

‘Nowcasting' of population alcohol-related harms using novel Bayesian timeseries methods and synthetic controls.

SG 16/17 235

Dr Frank de Vocht, Ms Cheryl McQuire, Professor Kate Tilling, Professor Matthew Hickman

APPENDIX

ACKNOWLEDGEMENTS

The authors would like to thank Peter Levell from the Institute for Fiscal Studies for providing the original data on time series of real alcohol duties used in Chapter 'Example: impact of alcohol taxation' of this report.

AUTHOR DETAILS

Dr Frank de Vocht. Senior Lecturer in Epidemiology and Public Health Research.
Population Health Sciences, Bristol Medical School
University of Bristol, Bristol, UK. Canynge Hall, 39 Whatley Road, Bristol, BS8 2PS.

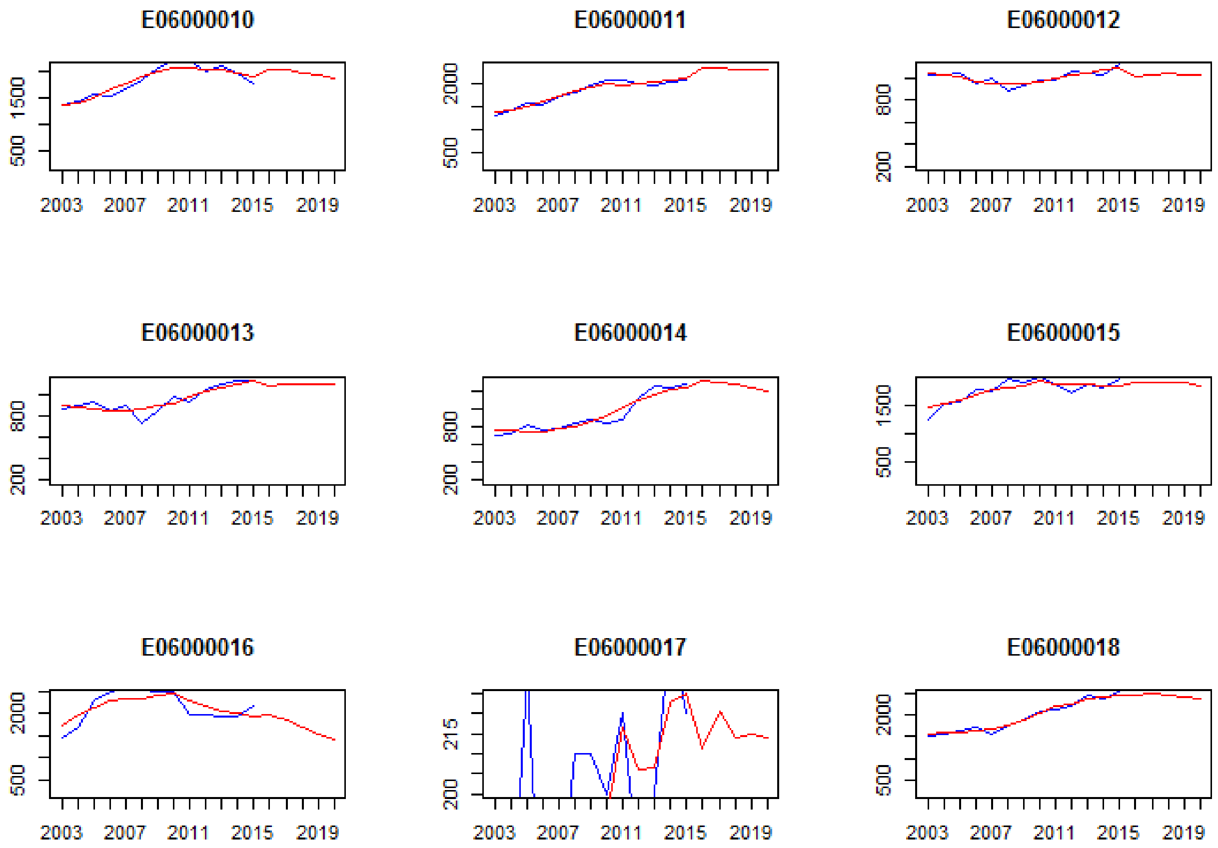
Email: frank.devocht@bristol.ac.uk.
Tel: +44.(0)117.928.7239
Postal address: Canynge Hall, 39 Whatley Road, Bristol, BS8 2PS.

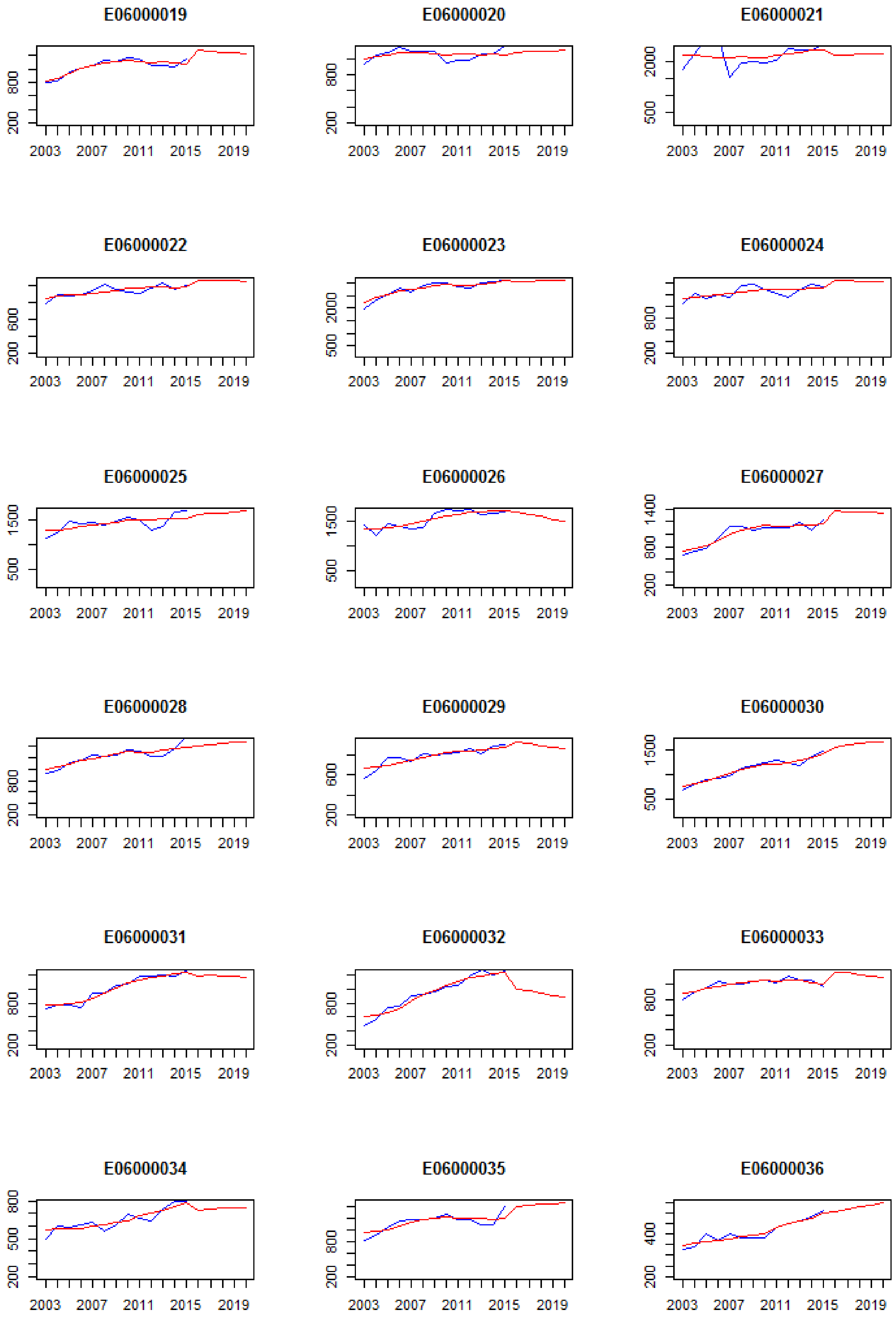
Ms Cheryl McQuire, Research Associate in Epidemiology and Alcohol Research
Population Health Sciences, Bristol Medical School
University of Bristol, Bristol, UK.
Email: cheryl.mcquire@bristol.ac.uk

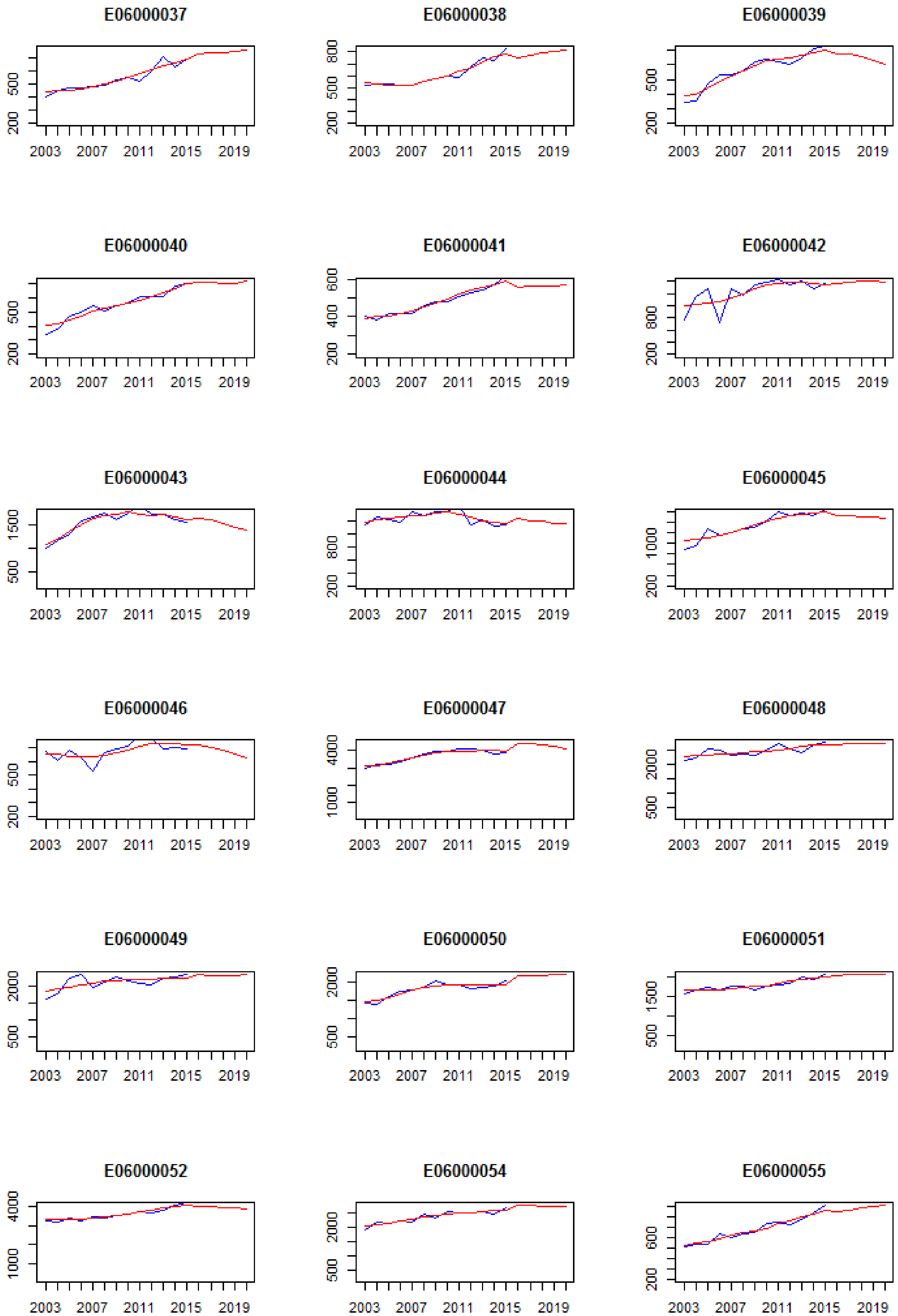
Professor Kate Tilling. Professor of Medical Statistics.
Population Health Sciences, Bristol Medical School
University of Bristol, Bristol, UK.
Email: kate.tilling@bristol.ac.uk

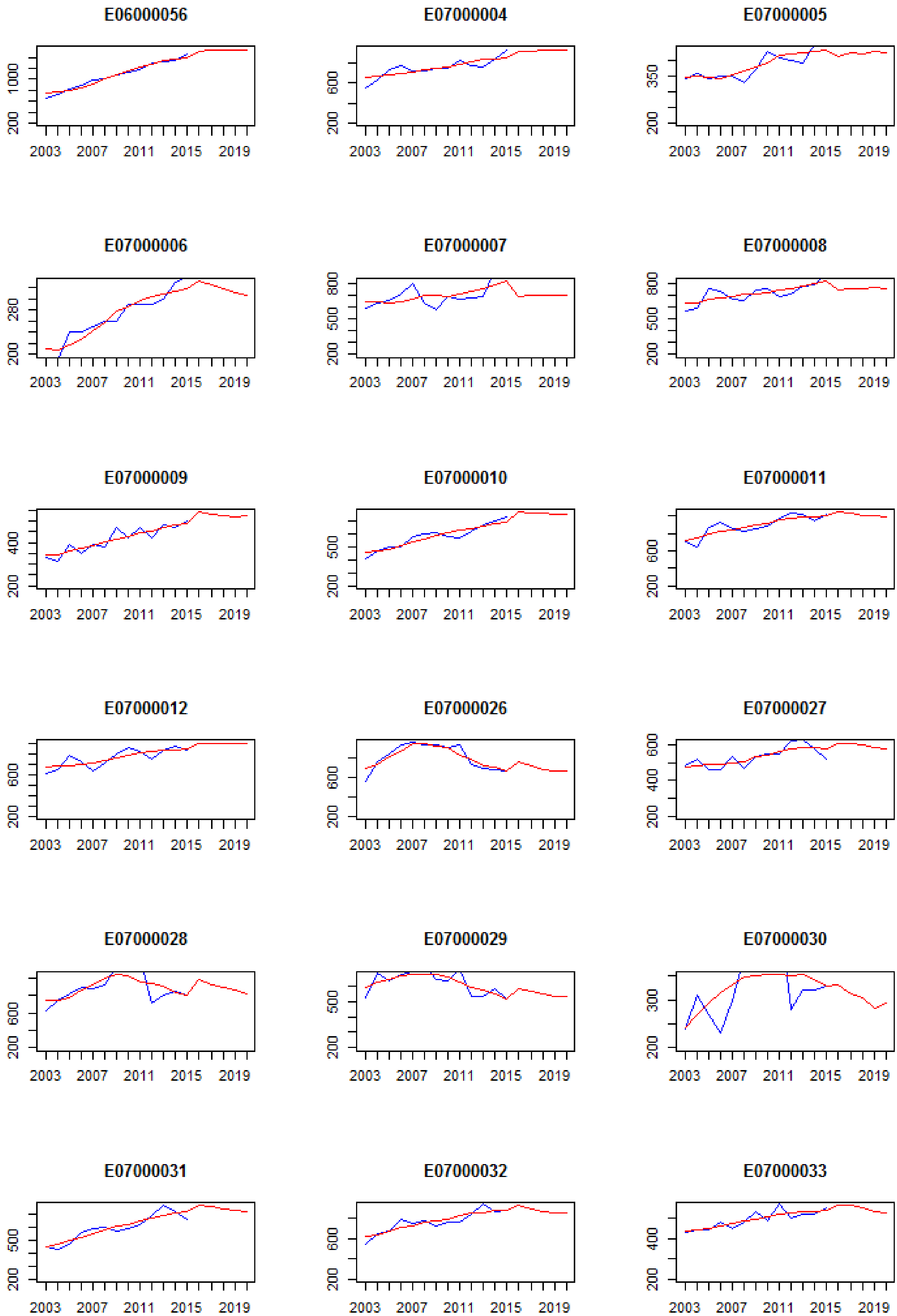
Professor Matthew Hickman. Professor of Public Health and Epidemiology.
Population Health Sciences, Bristol Medical School
University of Bristol, Bristol, UK.
Email: matthew.hickman@bristol.ac.uk

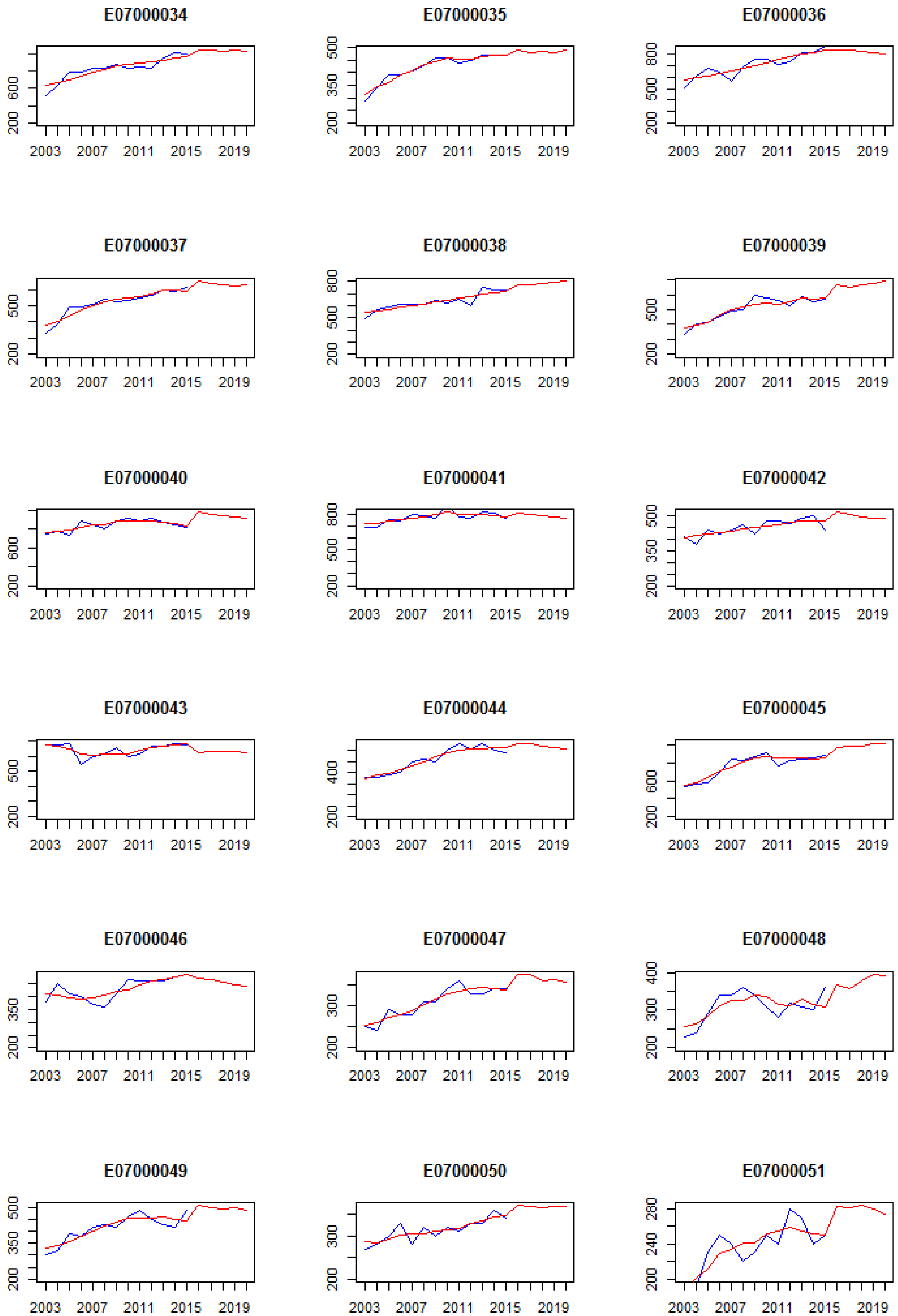
This report was funded by Alcohol Research UK. Alcohol Research UK is an independent charity working to reduce alcohol-related harm through ensuring policy and practice can be developed on the basis of reliable, research-based evidence.
www.alcoholresearchuk.org

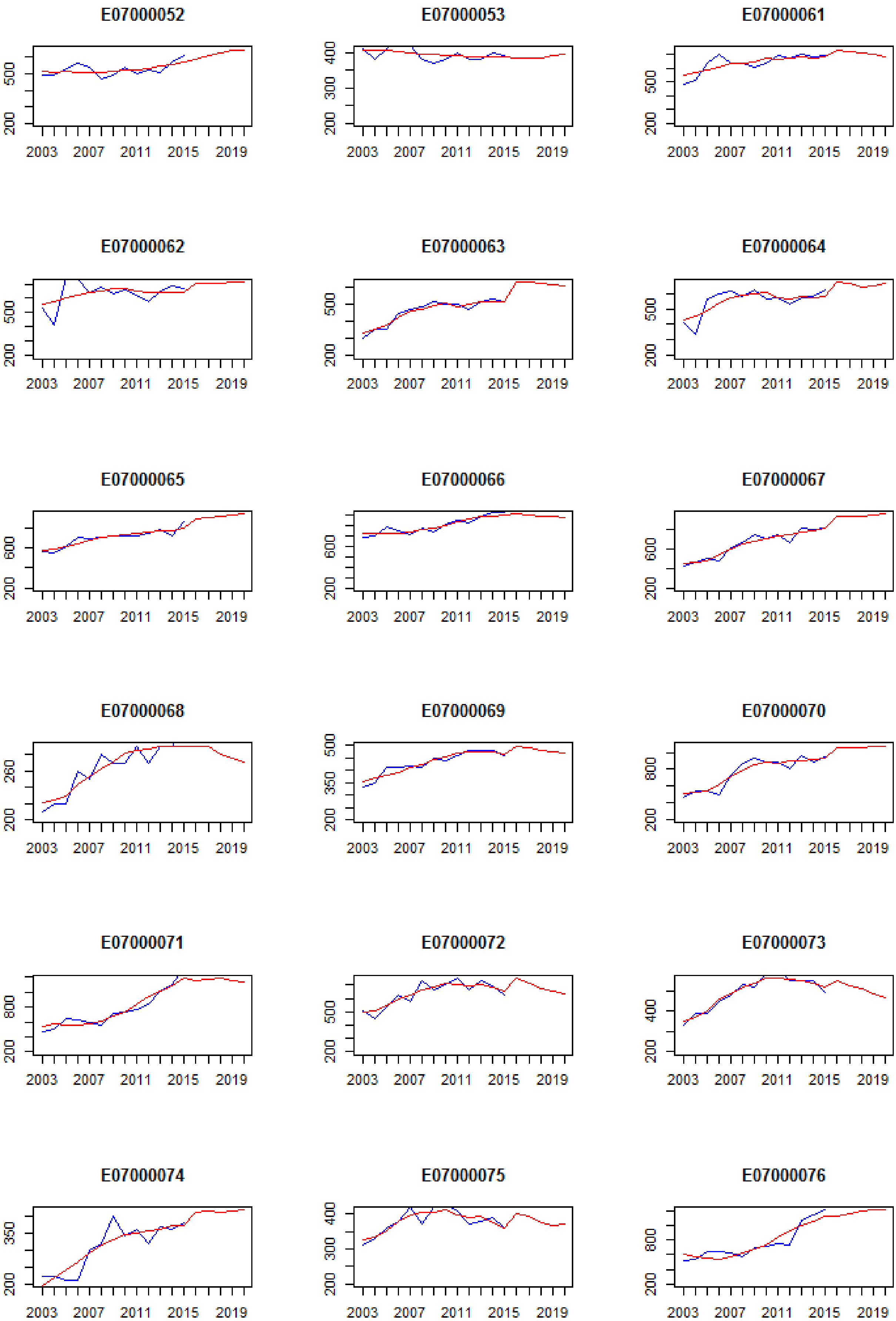

E06000010
E06000011
E06000012
E06000013
E06000014
E06000015
E06000016
E06000017
E06000018
2003 2007 2011 2015 2019

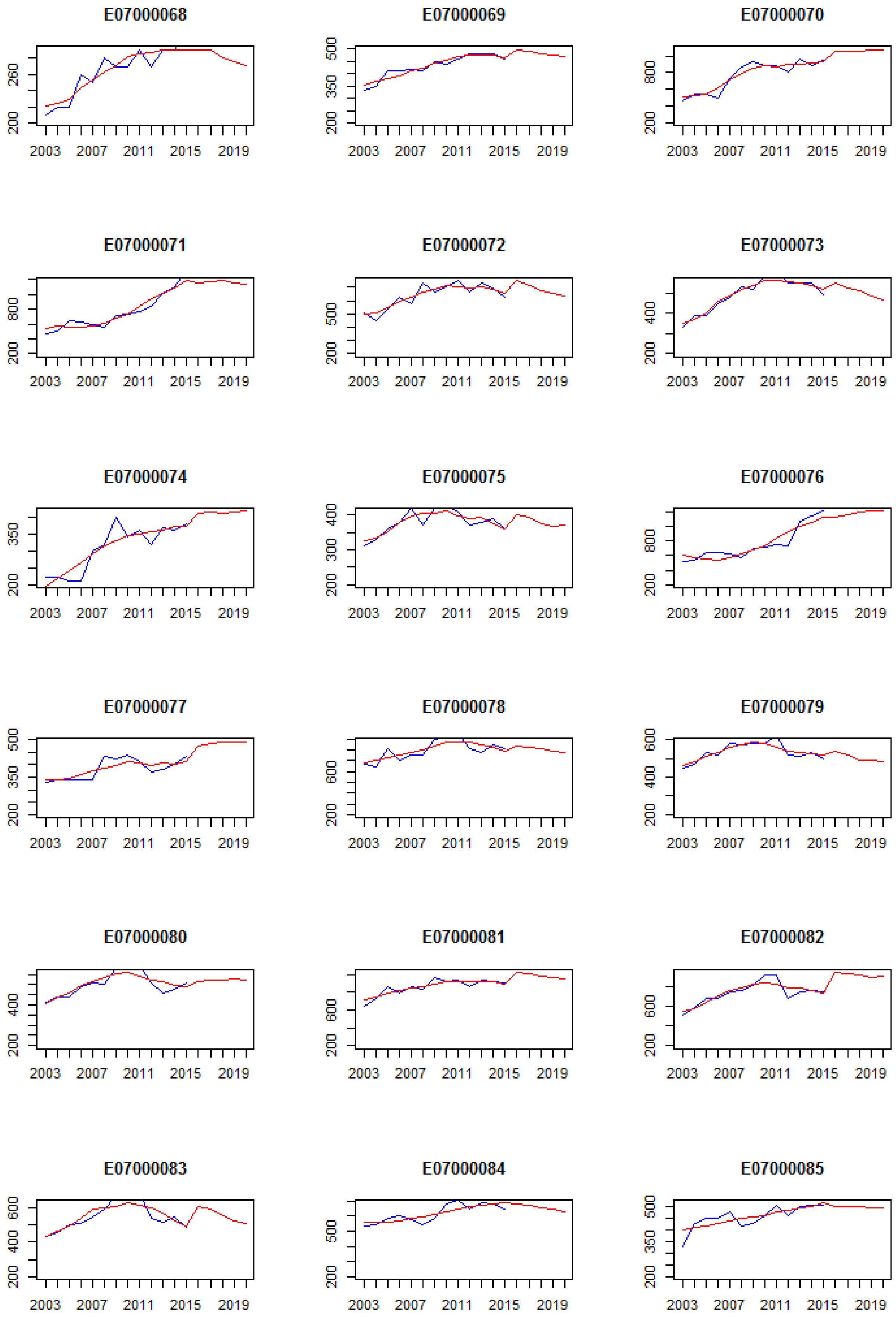

E06000019
E06000020
E06000021
E06000022
E06000023
E06000024
E06000025
E06000026
E06000027
E06000028
E06000029
E06000030
E06000031
E06000032
E06000033
E06000034
E06000035
E06000036

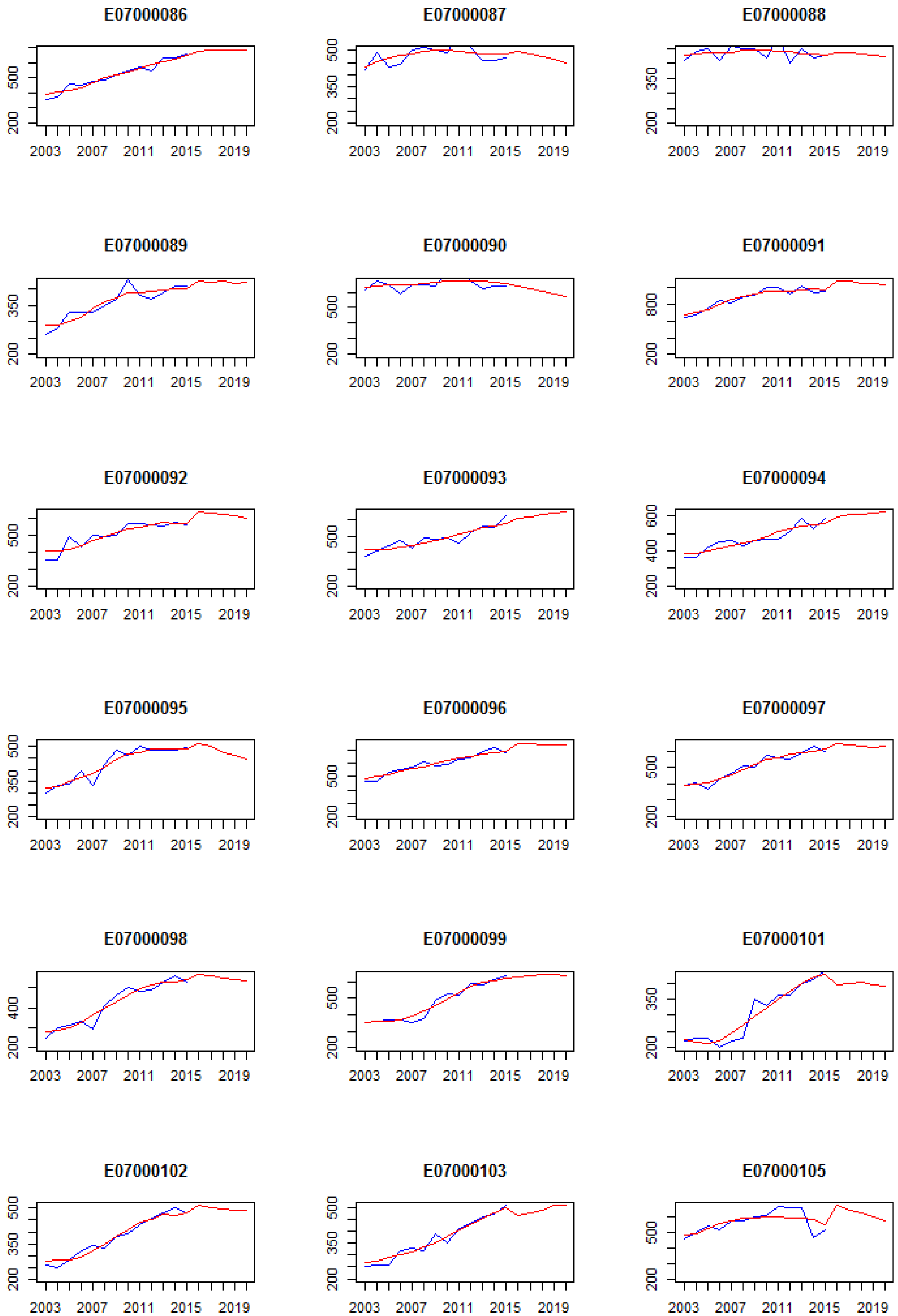

E06000037
200 500
2003 2007 2011 2015 2019
E06000038
200 500 800
2003 2007 2011 2015 2019
E06000039
200 500
2003 2007 2011 2015 2019
E06000040
200 500
2003 2007 2011 2015 2019
E06000041
200 400 600
2003 2007 2011 2015 2019
E06000042
200 800
2003 2007 2011 2015 2019
E06000043
500 1500
2003 2007 2011 2015 2019
E06000044
200 800
2003 2007 2011 2015 2019
E06000045
200 1000
2003 2007 2011 2015 2019
E06000046
200 500
2003 2007 2011 2015 2019
E06000047
1000 4000
2003 2007 2011 2015 2019
E06000048
500 2000
2003 2007 2011 2015 2019
E06000049
500 2000
2003 2007 2011 2015 2019
E06000050
500 2000
2003 2007 2011 2015 2019
E06000051
500 1500
2003 2007 2011 2015 2019
E06000052
1000 4000
2003 2007 2011 2015 2019
E06000054
500 2000
2003 2007 2011 2015 2019
E06000055
200 600
2003 2007 2011 2015 2019

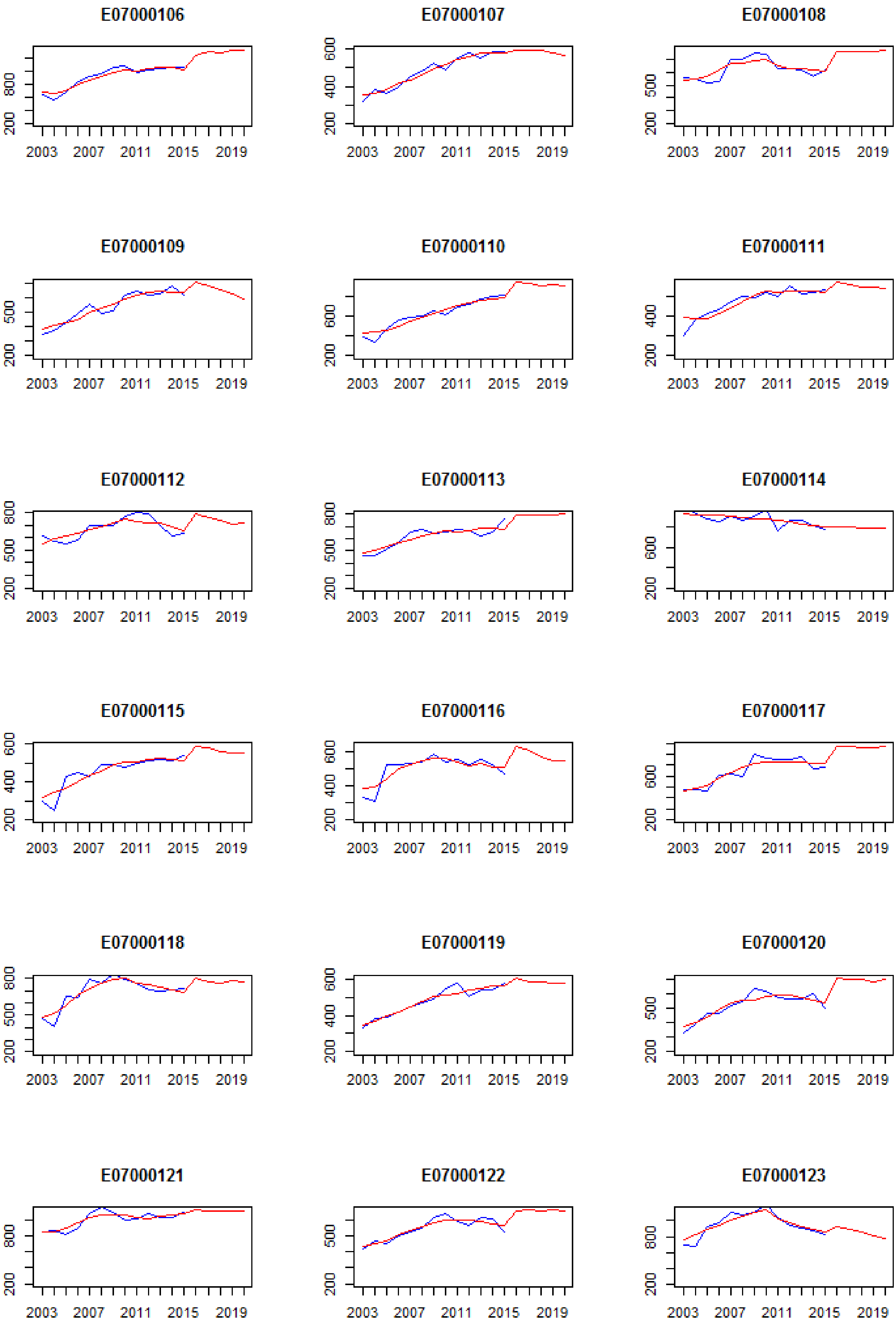

E06000056
E07000004
E07000005
E07000006
E07000007
E07000008
E07000009
E07000010
E07000011
E07000012
E07000026
E07000027
E07000028
E07000029
E07000030
E07000031
E07000032
E07000033
2003 2007 2011 2015 2019

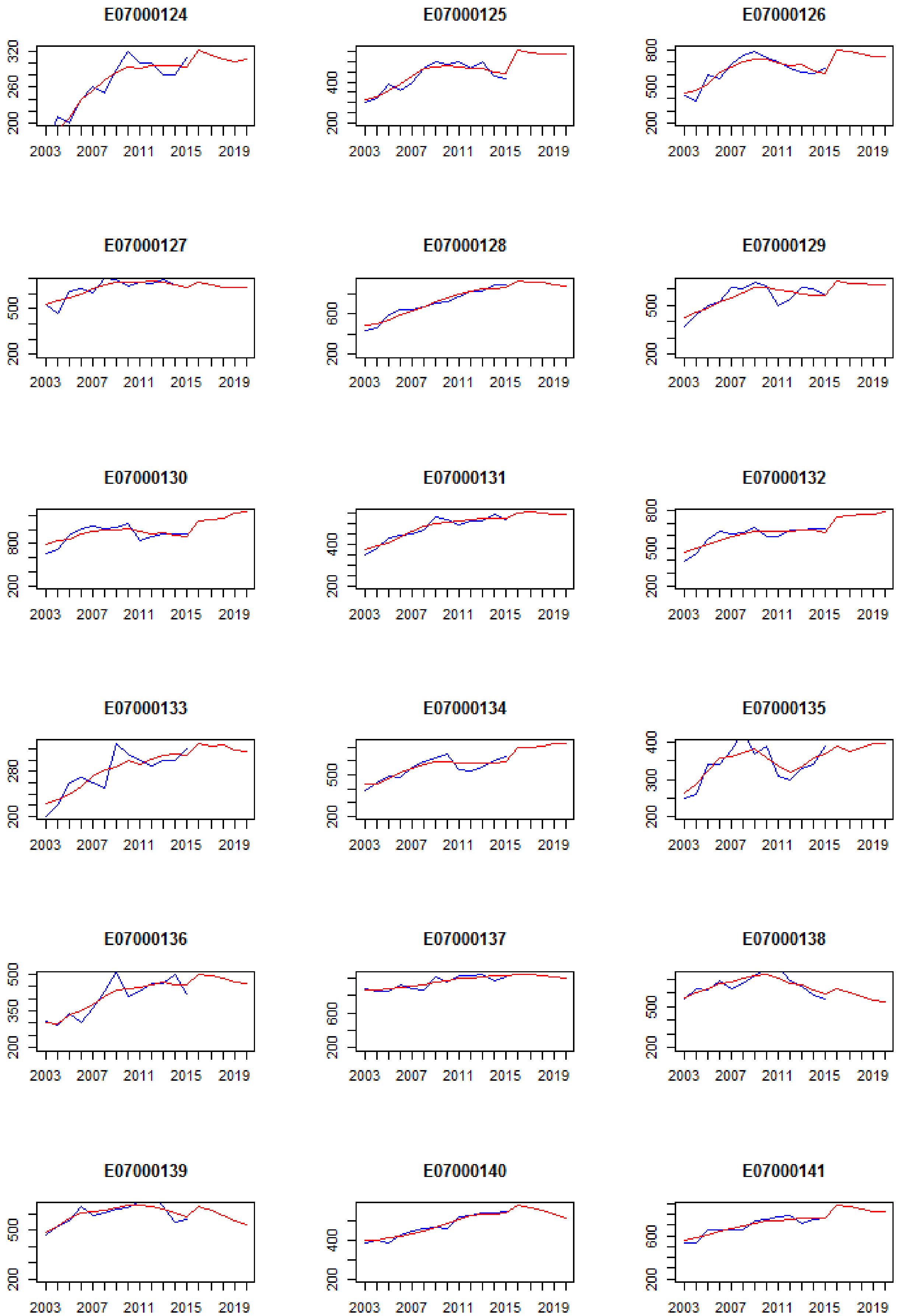

E07000034
E07000035
E07000036
E07000037
E07000038
E07000039
E07000040
E07000041
E07000042
E07000043
E07000044
E07000045
E07000046
E07000047
E07000048
E07000049
E07000050
E07000051
2003 2007 2011 2015 2019

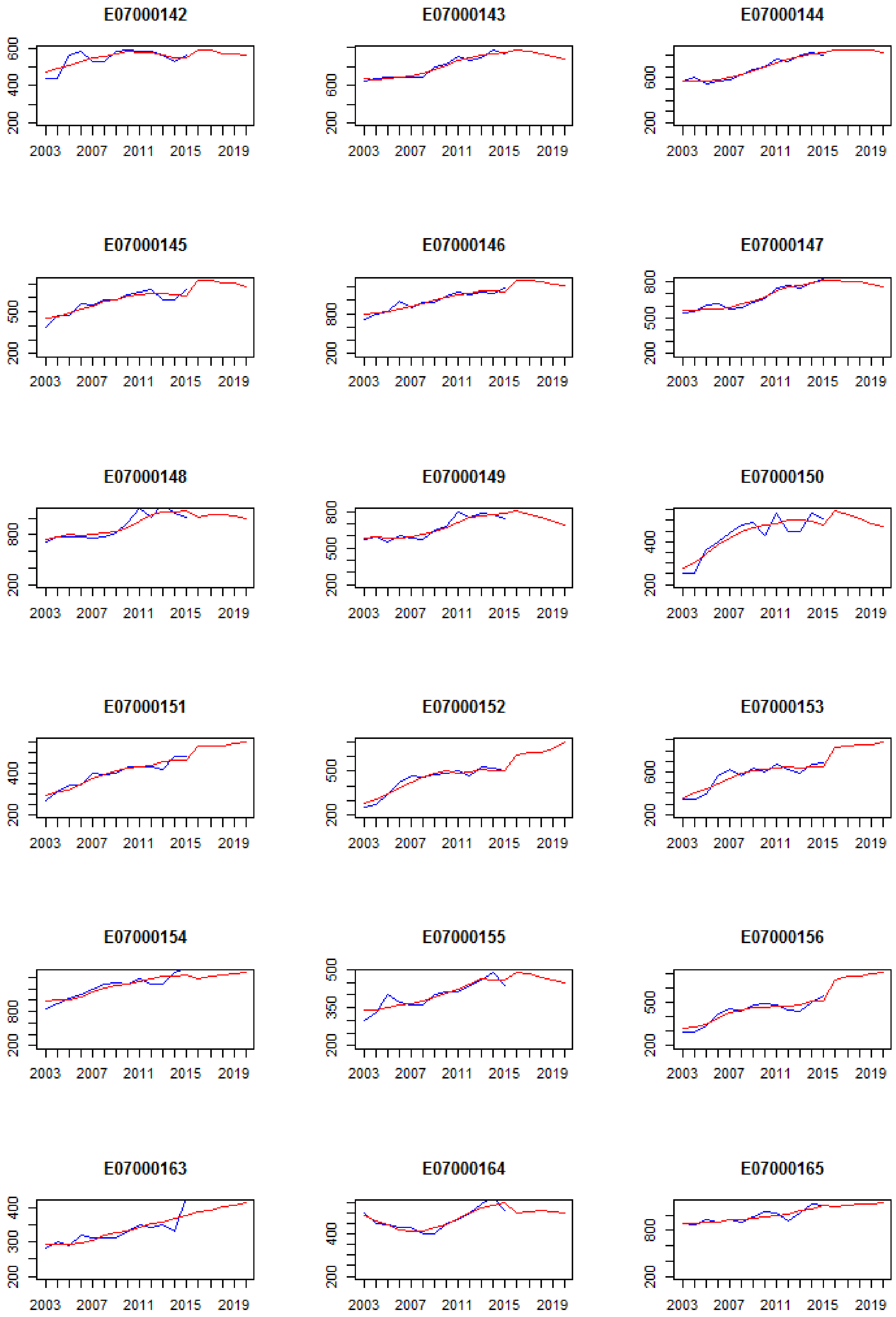

E07000052
E07000053
E07000061
E07000062
E07000063
E07000064
E07000065
E07000066
E07000067
E07000068
E07000069
E07000070
E07000071
E07000072
E07000073
E07000074
E07000075
E07000076
2003 2007 2011 2015 2019

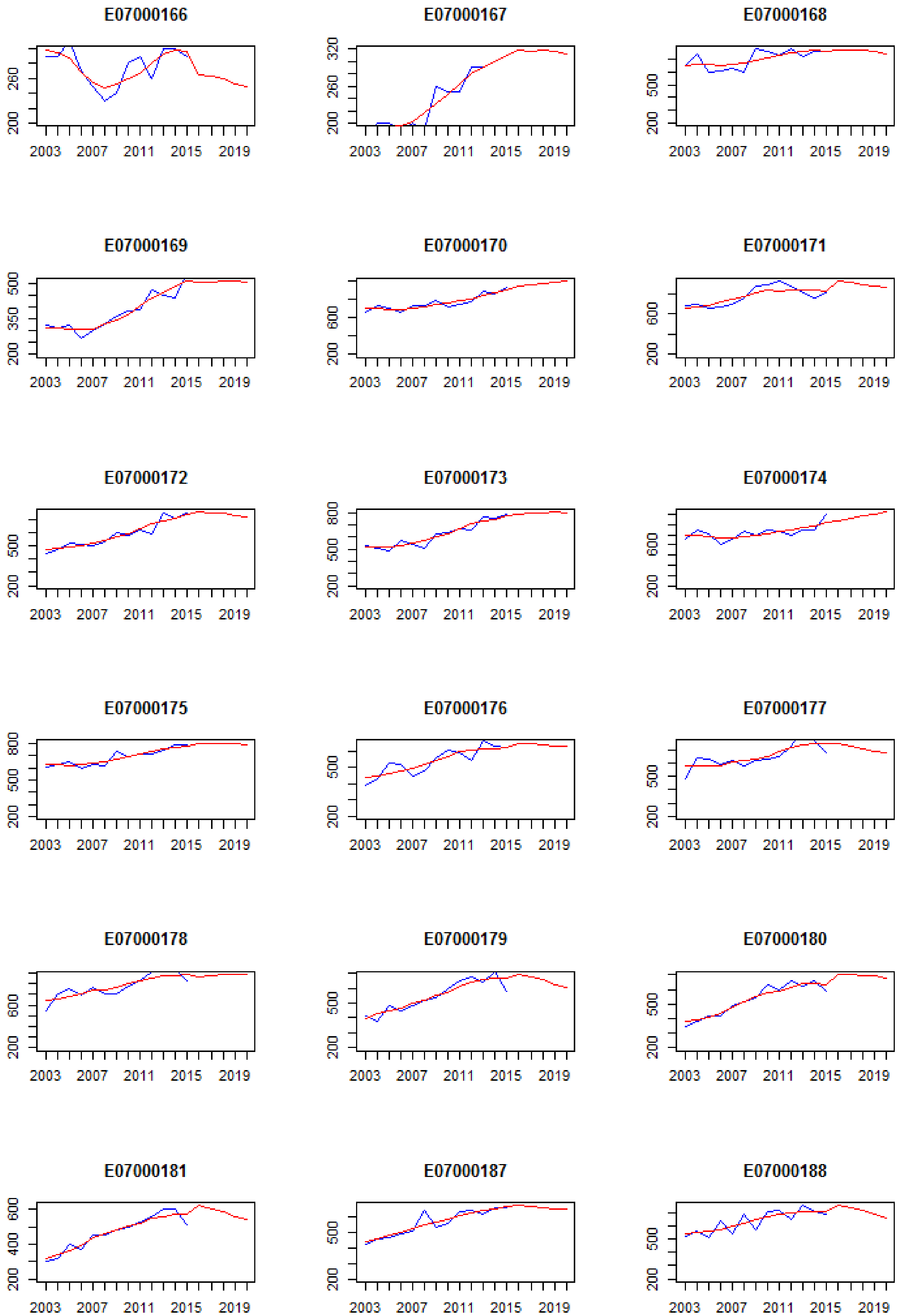

E07000068
E07000069
E07000070
E07000071
E07000072
E07000073
E07000074
E07000075
E07000076
E07000077
E07000078
E07000079
E07000080
E07000081
E07000082
E07000083
E07000084
E07000085

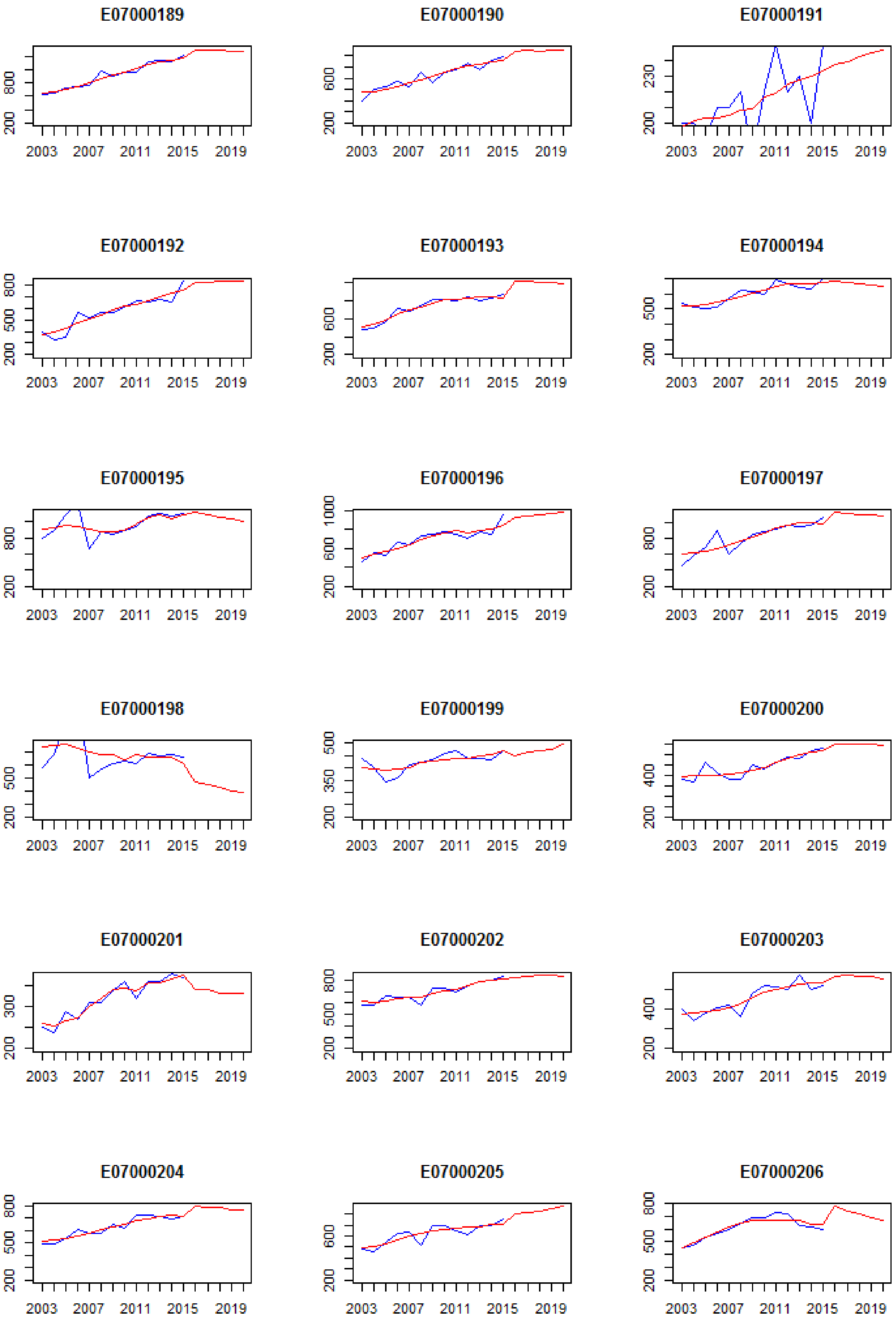

E07000086
200 500
2003 2007 2011 2015 2019
E07000087
200 350 500
2003 2007 2011 2015 2019
E07000088
200 350
2003 2007 2011 2015 2019
E07000089
200 350
2003 2007 2011 2015 2019
E07000090
200 500
2003 2007 2011 2015 2019
E07000091
200 800
2003 2007 2011 2015 2019
E07000092
200 500
2003 2007 2011 2015 2019
E07000093
200 500
2003 2007 2011 2015 2019
E07000094
200 400 600
2003 2007 2011 2015 2019
E07000095
200 350 500
2003 2007 2011 2015 2019
E07000096
200 500
2003 2007 2011 2015 2019
E07000097
200 500
2003 2007 2011 2015 2019
E07000098
200 400
2003 2007 2011 2015 2019
E07000099
200 500
2003 2007 2011 2015 2019
E07000101
200 350
2003 2007 2011 2015 2019
E07000102
200 350 500
2003 2007 2011 2015 2019
E07000103
200 350 500
2003 2007 2011 2015 2019
E07000105
200 500
2003 2007 2011 2015 2019

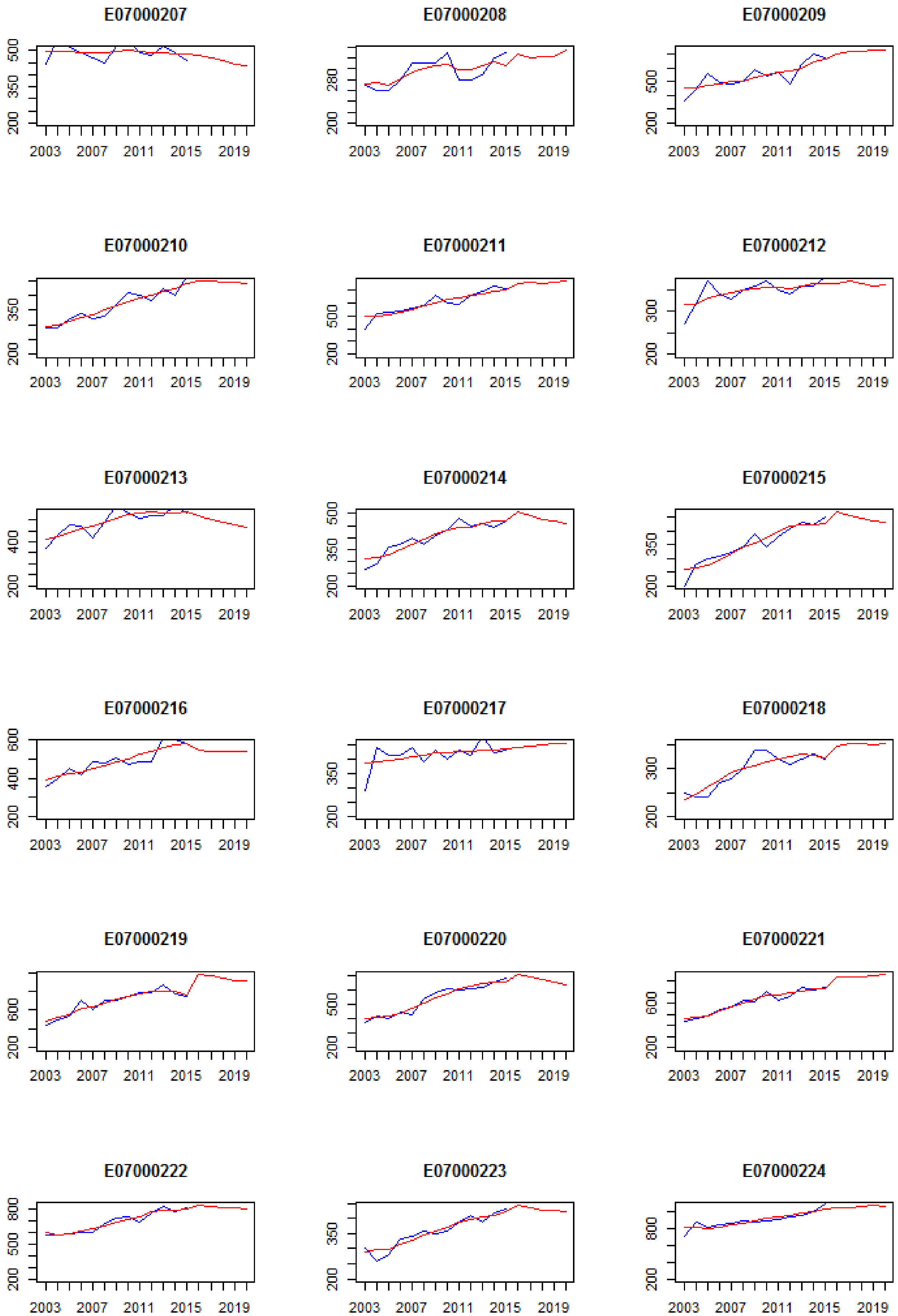

E07000106
200
800
2003 2007 2011 2015 2019
E07000107
200 400 600
2003 2007 2011 2015 2019
E07000108
200
500
2003 2007 2011 2015 2019
E07000109
200
500
2003 2007 2011 2015 2019
E07000110
200
600
2003 2007 2011 2015 2019
E07000111
200
400
2003 2007 2011 2015 2019
E07000112
200 500 800
2003 2007 2011 2015 2019
E07000113
200 500 800
2003 2007 2011 2015 2019
E07000114
200
600
2003 2007 2011 2015 2019
E07000115
200 400 600
2003 2007 2011 2015 2019
E07000116
200 400 600
2003 2007 2011 2015 2019
E07000117
200
600
2003 2007 2011 2015 2019
E07000118
200 500 800
2003 2007 2011 2015 2019
E07000119
200 400 600
2003 2007 2011 2015 2019
E07000120
200
500
2003 2007 2011 2015 2019
E07000121
200
800
2003 2007 2011 2015 2019
E07000122
200
500
2003 2007 2011 2015 2019
E07000123
200
800
2003 2007 2011 2015 2019

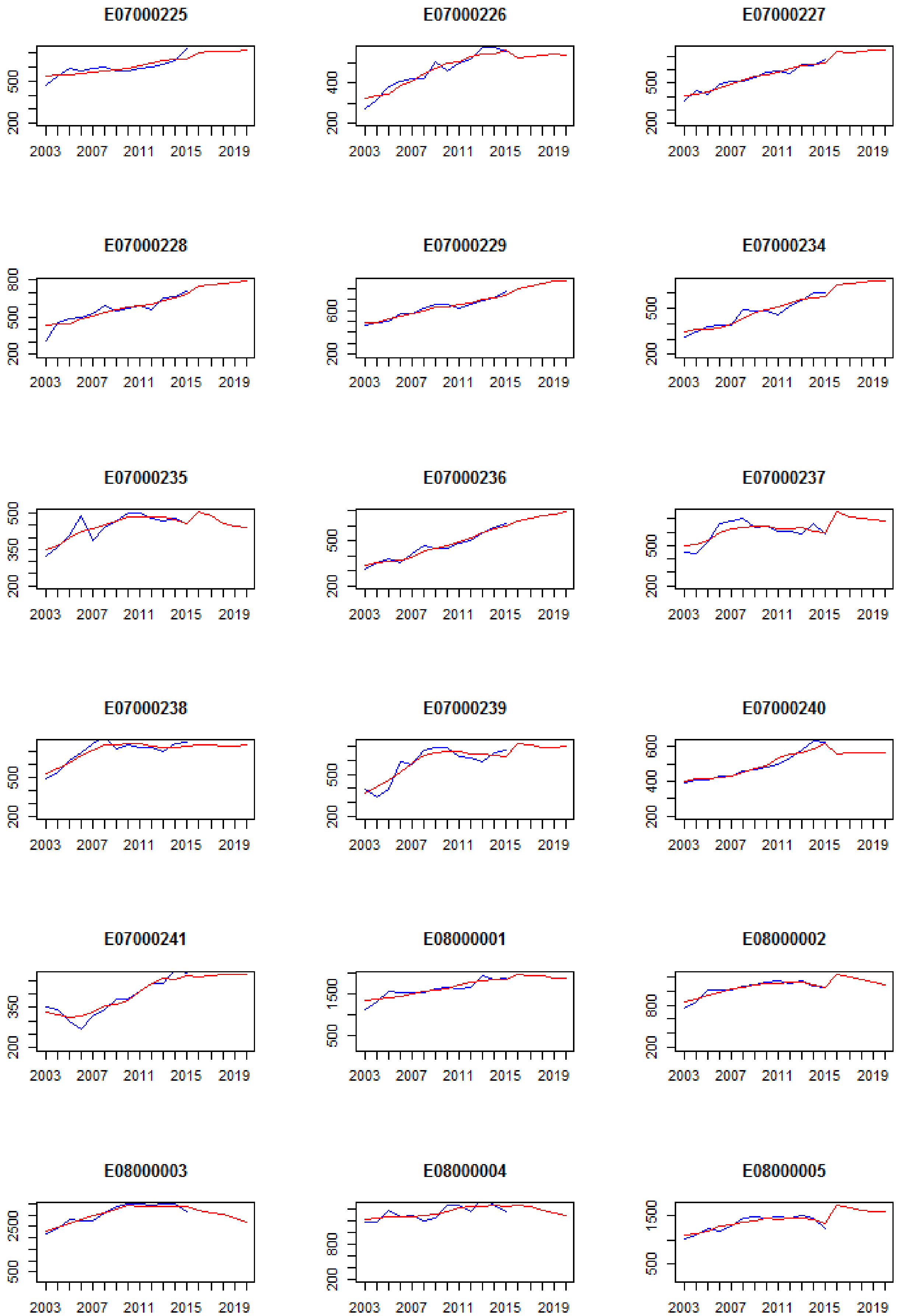

E07000124
E07000125
E07000126
E07000127
E07000128
E07000129
E07000130
E07000131
E07000132
E07000133
E07000134
E07000135
E07000136
E07000137
E07000138
E07000139
E07000140
E07000141
2003
2007
2011
2015
2019

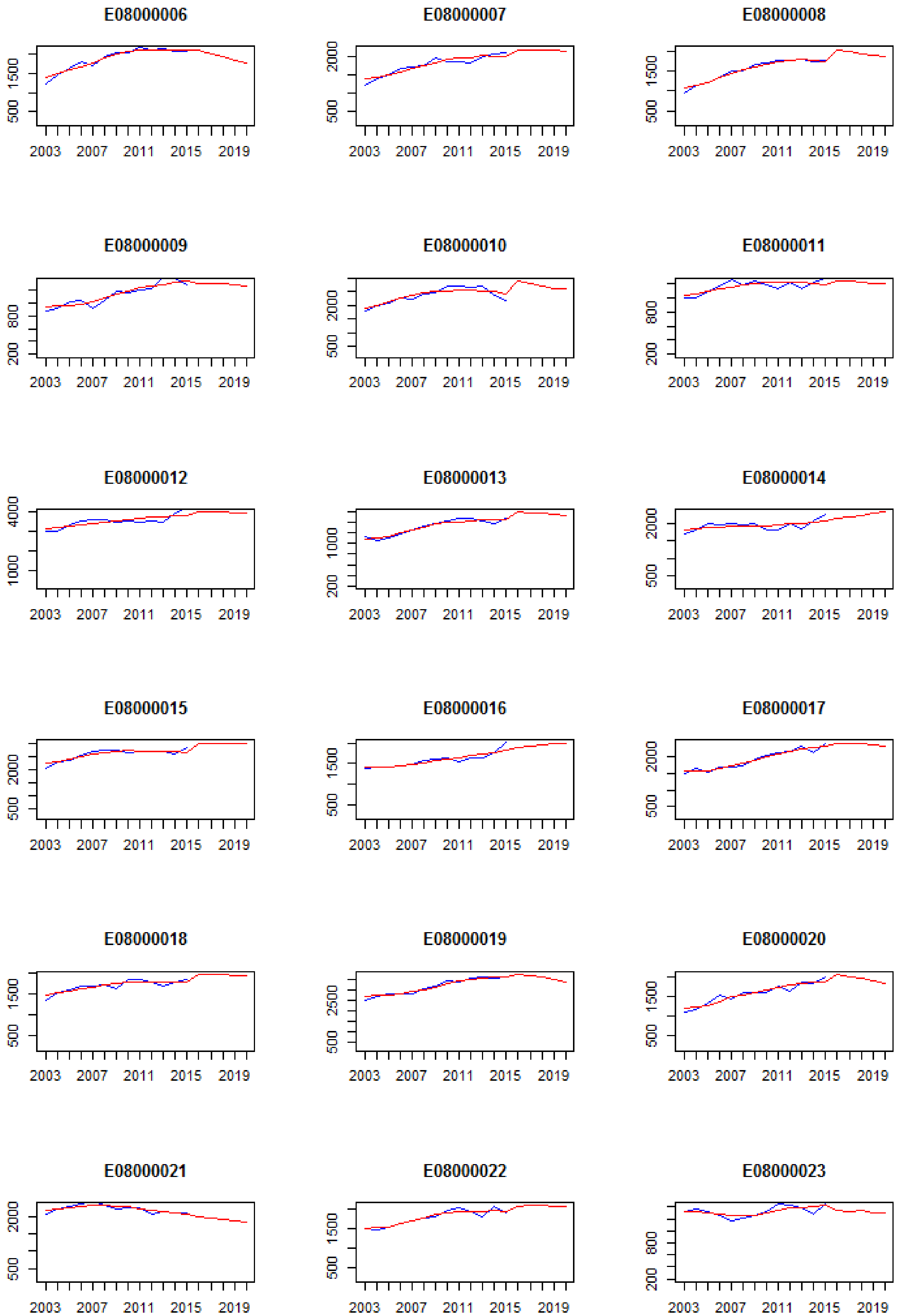

E07000142
E07000143
E07000144
E07000145
E07000146
E07000147
E07000148
E07000149
E07000150
E07000151
E07000152
E07000153
E07000154
E07000155
E07000156
E07000163
E07000164
E07000165
2003 2007 2011 2015 2019

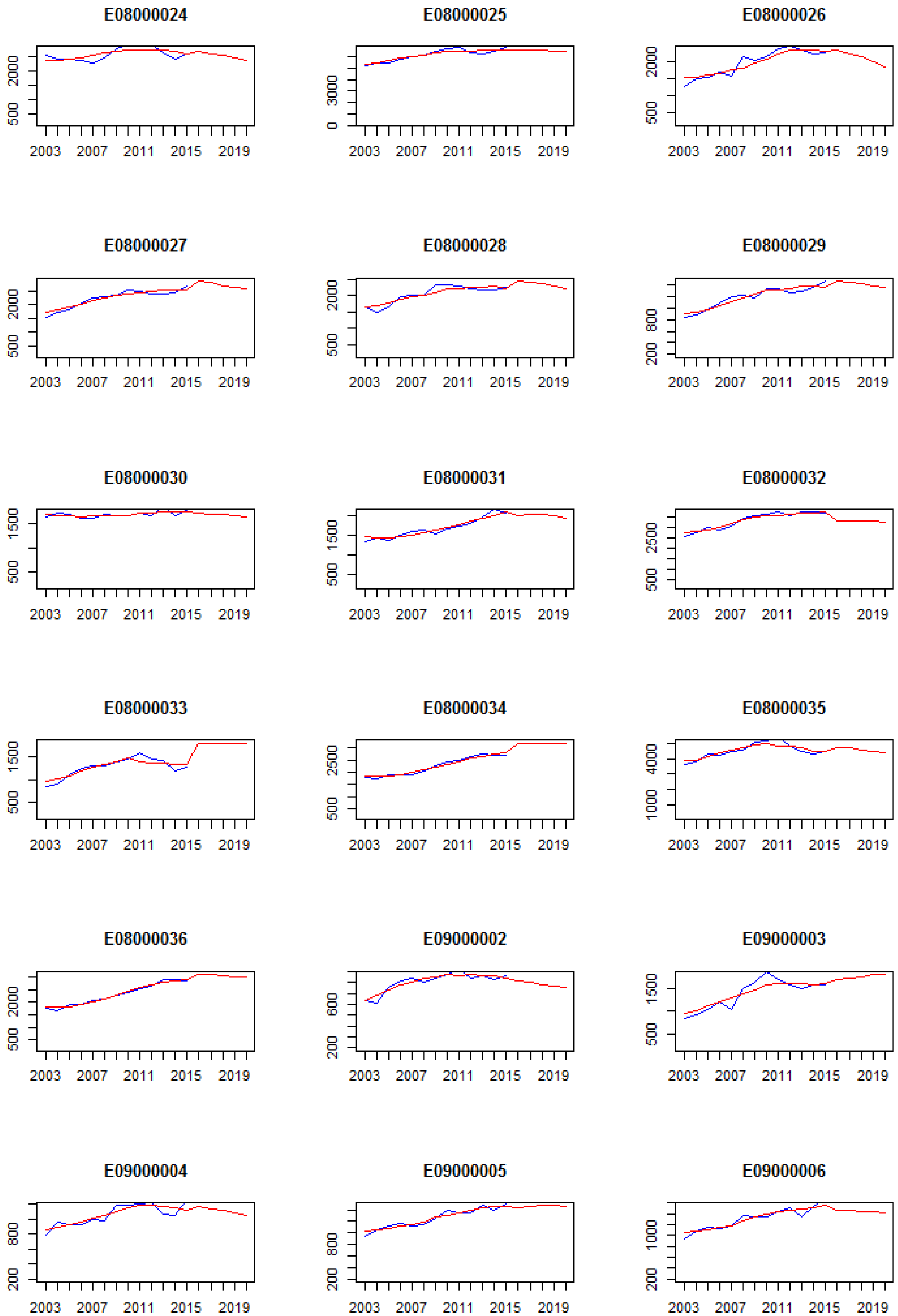

E07000166
200
260
E07000167
200
260
320
E07000168
200
500
E07000169
200
350
500
E07000170
200
600
E07000171
200
600
E07000172
200
500
E07000173
200
500
800
E07000174
200
600
E07000175
200
500
800
E07000176
200
500
E07000177
200
500
E07000178
200
600
E07000179
200
500
E07000180
200
500
E07000181
200
400
600
E07000187
200
500
E07000188
200
500
2003
2007
2011
2015
2019

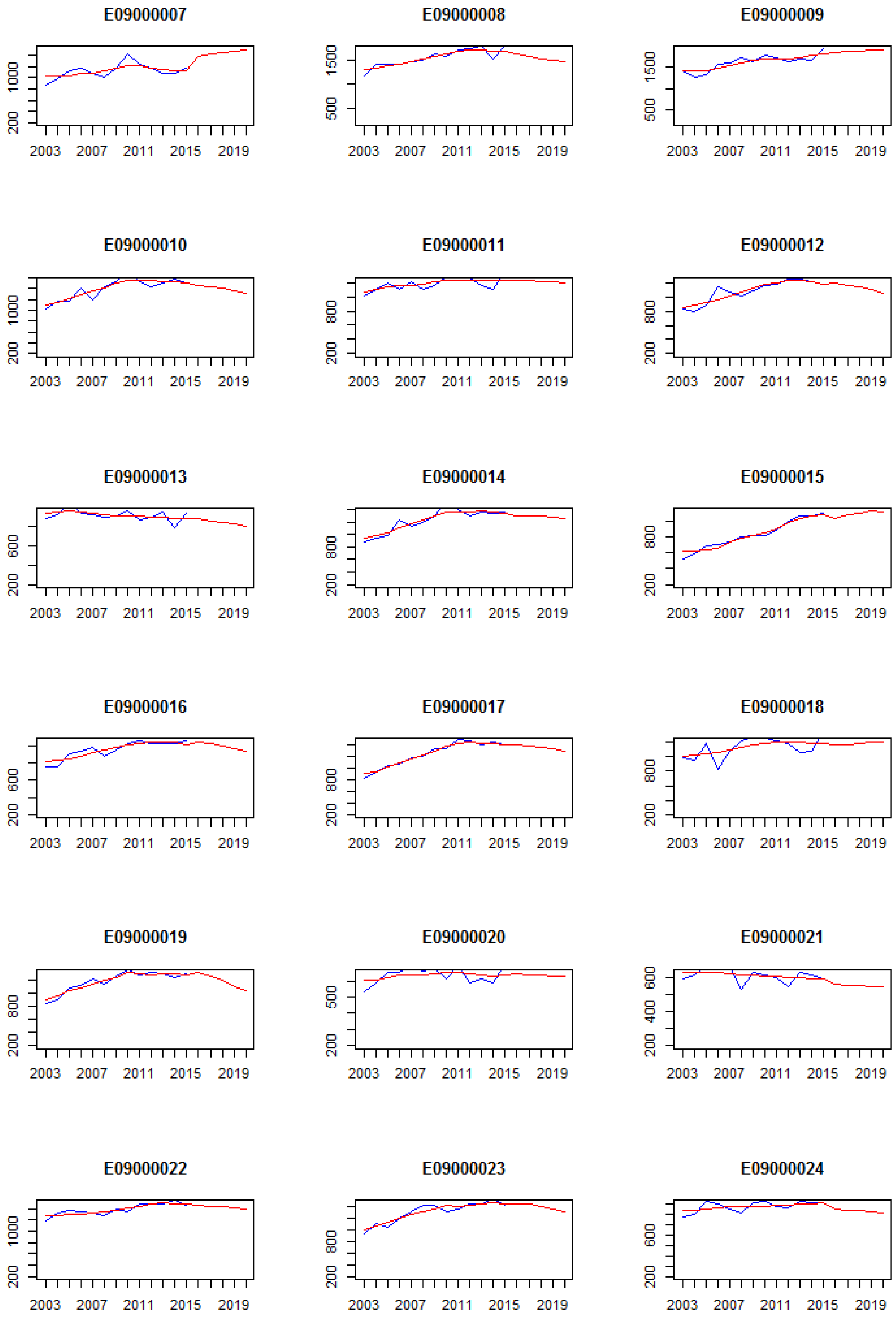

E07000189
E07000190
E07000191
E07000192
E07000193
E07000194
E07000195
E07000196
E07000197
E07000198
E07000199
E07000200
E07000201
E07000202
E07000203
E07000204
E07000205
E07000206
2003 2007 2011 2015 2019

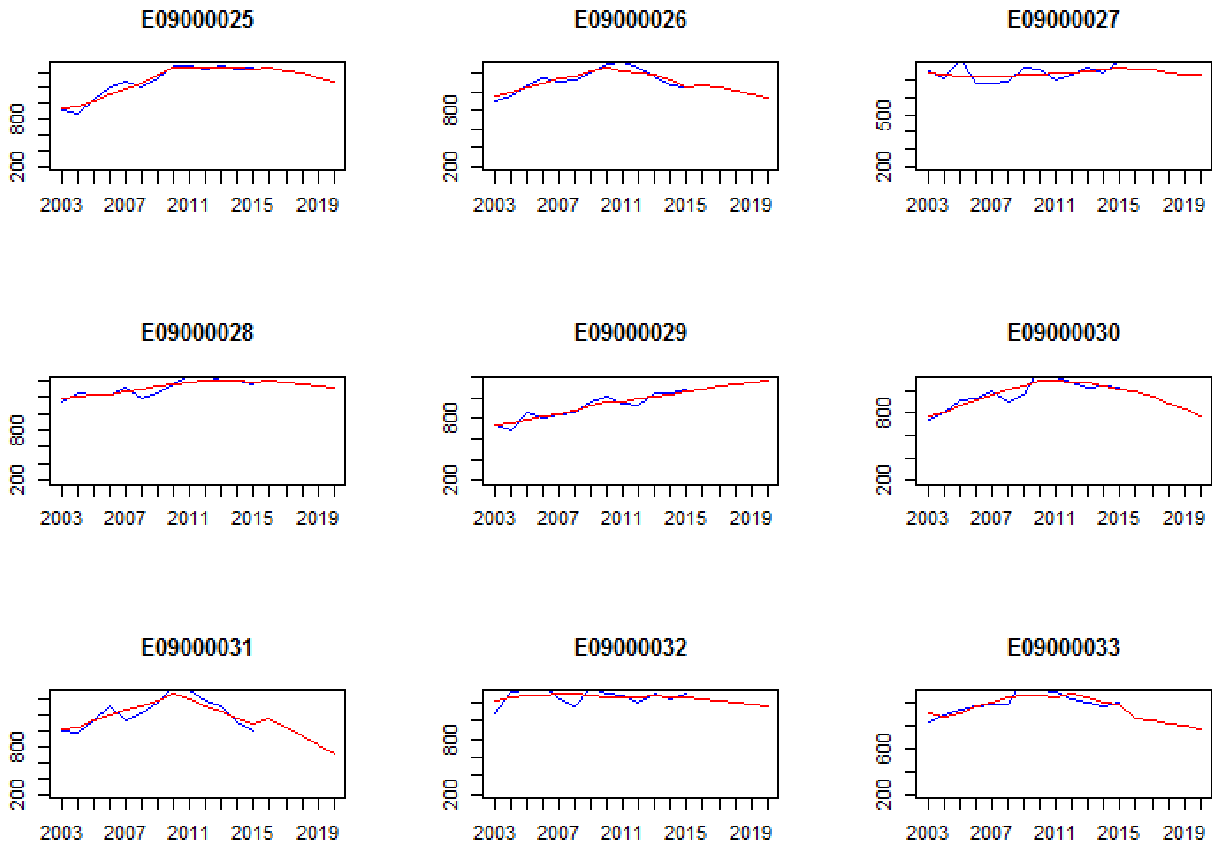

E07000207
200
350
500
E07000208
200
280
E07000209
200
500
E07000210
200
350
E07000211
200
500
E07000212
200
300
E07000213
200
400
E07000214
200
350
500
E07000215
200
350
E07000216
200
400
600
E07000217
200
350
E07000218
200
300
E07000219
200
600
E07000220
200
500
E07000221
200
600
E07000222
200
500
800
E07000223
200
350
E07000224
200
800
2003
2007
2011
2015
2019


E07000225
200 500
2003 2007 2011 2015 2019
E07000226
200 400
2003 2007 2011 2015 2019
E07000227
200 500
2003 2007 2011 2015 2019
E07000228
200 500 800
2003 2007 2011 2015 2019
E07000229
200 600
2003 2007 2011 2015 2019
E07000234
200 500
2003 2007 2011 2015 2019
E07000235
200 350 500
2003 2007 2011 2015 2019
E07000236
200 500
2003 2007 2011 2015 2019
E07000237
200 500
2003 2007 2011 2015 2019
E07000238
200 500
2003 2007 2011 2015 2019
E07000239
200 500
2003 2007 2011 2015 2019
E07000240
200 400 600
2003 2007 2011 2015 2019
E07000241
200 350
2003 2007 2011 2015 2019
E08000001
500 1500
2003 2007 2011 2015 2019
E08000002
200 800
2003 2007 2011 2015 2019
E08000003
500 2500
2003 2007 2011 2015 2019
E08000004
200 800
2003 2007 2011 2015 2019
E08000005
500 1500
2003 2007 2011 2015 2019


E08000006
E08000007
E08000008
E08000009
E08000010
E08000011
E08000012
E08000013
E08000014
E08000015
E08000016
E08000017
E08000018
E08000019
E08000020
E08000021
E08000022
E08000023


E08000024
500 2000
E08000025
0 3000
E08000026
500 2000
E08000027
500 2000
E08000028
500 2000
E08000029
200 800
E08000030
500 1500
E08000031
500 1500
E08000032
500 2500
E08000033
500 1500
E08000034
500 2500
E08000035
1000 4000
E08000036
500 2000
E09000002
200 600
E09000003
500 1500
E09000004
200 800
E09000005
200 800
E09000006
200 1000
2003 2007 2011 2015 2019


E09000007
200
1000
E09000008
500
1500
E09000009
500
1500
E09000010
200
1000
E09000011
200
800
E09000012
200
800
E09000013
200
600
E09000014
200
800
E09000015
200
800
E09000016
200
600
E09000017
200
800
E09000018
200
800
E09000019
200
800
E09000020
200
500
E09000021
200
400
600
E09000022
200
1000
E09000023
200
800
E09000024
200
600
2003
2007
2011
2015
2019


E09000025
200
800
2003 2007 2011 2015 2019
E09000026
200
800
2003 2007 2011 2015 2019
E09000027
200
500
2003 2007 2011 2015 2019
E09000028
200
800
2003 2007 2011 2015 2019
E09000029
200
800
2003 2007 2011 2015 2019
E09000030
200
800
2003 2007 2011 2015 2019
E09000031
200
800
2003 2007 2011 2015 2019
E09000032
200
800
2003 2007 2011 2015 2019
E09000033
200
600
2003 2007 2011 2015 2019

EXAMPLE SCRIPT AND BAYESIAN STRUCTURAL TIME-SERIES

Find below the script to run the 5-yr nowcasting model (2002/3 – 2015/16), including assessment of error, for a randomly selected lower-tier local area. The below script is not identical to the ones used in this report but, for ease of use, uses all default settings of the Bayesian structural time series method. More information can be found at: https://google.github.io/CausalImpact/CausalImpact.html

Note that for this analysis to run, you require a copy of *R* statistical software with the 'Causal Inference' R package installed.

```
# Creation of the dataset
hosp.adm<-c(210,220,220,260,250,280,270,270,290,270,290,290,310)
age.15<-c(13549,13608,13689,13654,13631,13686,13582,13486,13468,13501,13650,13895,14114)
age.1624<-c(6073,6234,6403,6660,6921,7033,7173,7301,7224,7207,7518,7319,7213)
age.2534<-c(8092,8049,7881,7829,7778,7840,7953,8104,8168,8247,8600,8769,8803)
age.3544<-c(11036,11112,11183,10967,10886,10715,10536,10270,10013,9924,9945,9833,9765)
age.4554<-c(9291,9517,9755,10010,10357,10727,11073,11393,11552,11720,11860,11891,11819)
age.5564<-c(8186,8173,8110,8391,8590,8790,8894,9044,8779,8652,8618,8657,8853)
age.6574<-c(6705,6638,6537,6419,6531,6644,6805,6921,7318,7566,7638,7740,7911)
age.75<-c(6277,6414,6563,6714,6792,6938,7094,7238,7384,7508,7631,7678,7656)
data<-cbind(hosp.adm,age.15,age.1624,age.2534,age.3544,age.4554,age.5564,age.6574,age.75)

#Calling the statistical package
library(CausalImpact)

# define the model period and the nowcasting period (note that you can change these for different
# nowcasting periods and run the model
length.pre<-8          #2002/3 to 2009/10
length.post<-5         #2010/11 to 2015/16
impact<-(CausalImpact(data, pre.period,post.period))

#obtain output
plot(impact)
summary(impact)
```