ESCoE — Economic Statistics Centre of Excellence

A COLLABORATION WITH

Office for National Statistics

Nowcasting Unpaid Production Activity and Quality Adjustments in Public Service Productivity

George Kapetanios and Fotis Papailias

ESCoE Discussion Paper 2023-20

November 2023

ISSN 2515-4664

DISCUSSION PAPER

Nowcasting Unpaid Production Activity and Quality Adjustments in Public Service Productivity
George Kapetanios and Fotis Papailias
ESCoE Discussion Paper No. 2023-20
November 2023

Abstract

Economists are so focused on key macroeconomic variables, such as the GDP, Inflation, Unemployment, Industrial Production, Retail Sales, etc., who often omit or skip variables which are more qualitative in nature but of key importance for social conditions and living. These variables include the Unpaid Household Production Activity as well as Quality Adjustment for Public Service Productivity. This research attempts to provide a generic methodological framework for a dynamic nowcasting of such variables and assist relevant economists and statisticians in tracking of these low frequency qualitative variables.

Keywords: Nowcasting, Factor Models, Penalised Regression, Neural Networks, Support Vector Regression

JEL classification: C53, E37

Fotis Papailias, King's College London
fotis.papailias@kcl.ac.uk

Published by:
Economic Statistics Centre of Excellence
King's College London
Strand
London
WC2R 2LS
United Kingdom
www.escoe.ac.uk

Nowcasting Unpaid Production Activity and Quality Adjustments in Public Service Productivity

George Kapetanios Fotis Papailias*

King's Business School, DAFM

Economic Statistics Centre of Excellence

November, 2023

Abstract

Economists are so focused on key macroeconomic variables, such as the GDP, Inflation, Unemployment, Industrial Production, Retail Sales, etc., who often omit or skip variables which are more qualitative in nature but of key importance for social conditions and living. These variables include the Unpaid Household Production Activity as well as Quality Adjustment for Public Service Productivity. This research attempts to provide a generic methodological framework for a dynamic nowcasting of such variables and assist relevant economists and statisticians in tracking of these low frequency qualitative variables.

Keywords: Nowcasting, Factor Models, Penalised Regression, Neural Networks, Support Vector Regression.

JEL Codes: C53, E37.

* Corresponding author. King's Business School, King's College London, Bush House, 30 Aldwych, London, WC2B 4BG, UK. Knot Analytics Ltd. E-mail: fotis.papailias@kcl.ac.uk

This research has been funded by the Office for National Statistics as part of the research programme of the Economic Statistics Centre of Excellence (ESCoE).

Contents

1 Introduction

Economists are often obsessed with key economic indicators that mainly track the balance of payments, the business and consumer tendency, the financial markets, the industry, the trade, the labour market, the level of prices and the national accounts.[1] It cannot be argued that this data is of utmost importance to track the economy in real time. However, variables which are more "qualitative" in nature but carry significant importance for social conditions and living are ignored or skipped. These indicators include, among others, the Unpaid Household Production Activity (UHP) as well as Quality Adjustment for Public Service Productivity (PSP) variables.

The Office for National Statistics (ONS) in the UK has been paying considerable efforts to accurately measure such variables which are important aspects of people's lives but are largely missing from regular economic statistics such as the GDP. The UHP variables include childcare, adult care, household services, as well as unpaid nutrition, transport, laundry and volunteering services.[2] In PSP, quality adjustments are currently applied to five service areas: healthcare, education, children's social care, adult social care, and public order and safety.[3]

This research provides a generic methodological framework to nowcast these qualitative variables by employing a large panel of potential predictors and models based on state-of-the-art machine learning approaches. To the best of our knowledge, this is the first paper that attempts to nowcast this type of economic variables. The empirical results are encouraging indicating that there exist models which can outperform the standard benchmarks which mainly consist of univariate autoregressive models. Following the recent empirical literature, we consider models from the following methodologies: best subset selection, sparse regressions, factor extraction, random forests, neural networks and support vector regression; see Kapetanios and Papailias (2021, 2022a, 2022b).

The rest of this paper is structured as follows. Section 2 briefly describes the datasets as well as data limitations. Section 3 provides the details of the various

[1]See here for OECD's list of main economic indicators.

[2]Methodological details on the measurement of such variabes can be found here and here.

[3]More details can be found here and here.

econometric methodologies. Section 4 discusses the out-of-sample nowcasting exercise design. Section 5 briefly discusses the empirical results. Finally, Section 6 offers the concluding remarks.

2 Data Description and Limitations

2.1 Targets

As discussed in the Introduction, our main interest in this research is to nowcast the UHP and PSP variables. Starting with the UHP variables we have 15 targets in total; 8 measure market-equivalent output and 7 measure full market-equivalent intermediate consumption. These are annual variables with a sample of *only* 12 observations (from 2005 to 2016). The PSP variables include most of the components for education, most of the components for public order and safety and the final quality adjusted indices for education, adults social care, children's social care and public order and saftery Table 1 provides the description of all target variables.

These variables are transformed to stationarity using the growth rates (calculates as the first difference of their natural logarithm). However, to measure the performance of each model, nowcasts are translated back to levels. Tables 2 and 3 provide the descriptive statistics for the log growth of the target variables for the UHP and PSP respectively.

An obvious limitation is the very short history and sample size of the target variables. First, a short sample does not allow for efficient estimation. Second, short history does not allow enough periods for the out-of-sample cross validation.[4] The suggested models, as described later, are based in state-of-the-art machine learning methodologies which do not suffer -in principle- by the curse of dimensionality. However, all results must be examined with caution.

[4]For the UHP targets we leave the last 4 years as out-of-sample (2013, 2014, 2015, 2016) which allows only 7 observations to be used in the first estimation (in log growth). Similarly, for the PSP targets our out-of-sample (2016, 2017, 2018, 2019) allows for 10 observations to be used in the first estimation.

2.2 Set of Potential Predictors

In an effort to nowcast the previously discussed targets, we organised a large panel of *potential* predictors which include variables from many categories such as energy consumption and carbon emissions (electricity, gas, coal, etc.), demographics (births, deaths), household statistics, migration statistics, climate change variables, human development indices, social and political indicators, inequality measures, freedom indices, corruption and rule of law indices, government budget, income statistics, current account statistics, capital account variables, balance sheet indicators, the GDP and its components, foreign trade, real estate, policy uncertainty indicators, international reserves, employment, unemployment and job vacancy statistics, monetary statistics, consumer and producer prices, business surveys, consumer surveys and leading indicators, retail trade, commodity prices, equity indices, sovereign bond yields and corporate fixed-income statistics, and forex and swap rates. Variables are expressed in annual, quarterly, monthly and weekly-turned-to-monthly frequencies.[5]

The *initial* set of variables is described in the accompanying .xlsx file. However, it is important to highlight that the *final* set of variables used varies from case to case and is subject to the following filtering operations.

1. First, each variable is checked for seasonalities. If seasonalities are detected they are removed using the Cleveland et al. (1990) STL decomposition method.

2. Then, quarterly and monthly (including weekly-turned-to-monthly) variables are translated to annual using the annual averages (a simplified Bridge equations approach).

3. Any missing values due to publication lags are exrapolated using the exponential average of the past 2 years.

4. Then, variables are transformed to stationarity (the accompanying .xlsx file provides the transformation codes for each variable).

5. Outliers are removed (if required).

[5]Weekly variables are translated to monthly equivalents taking the monthly averages.

6. Collinear variables are removed; we define “colinear” variables as an pair of variables with an absolute correlation greater than 85%. Even though most of the state-of-the-art machine learning methodologies we employ are able to handle colinearities, removing them beforehand ensures a more efficient estimation.

Finally, it is important to note that we distinguish two cases for this large dataset: (i) in the first case we use unconditionally all the (filtered) variables, (ii) in the second case, we do a manual pre-selection of a smaller subset of these variables.[6] In principle, the first larger dataset should provide better estimates as models are able to “learn” from a larger universe of variables. However, there is a line of research in the literature which provides evidence that this is not always the case in empirical applications; see, among others, Camba-Mendez et al. (2001), Boivin and Ng (2006), Caggiano et al. (2011) and Camba-Mendez et al. (2015).

3 Econometric Setup

Let y_t, $t = 1, ..., T$, be the target variable and $x_t = (x_{1t}, ..., x_{Nt})'$ be the main set of potential predictors. It is not necessary to assume a particular data generating process for y_t but simply posit the existence of a representation of the form:

$$y_t = a + f(x_{1t}, ..., x_{Nt}) + u_t. \tag{1}$$

While the potential nonlinearity in Equation (1) might, in principle, be worth exploring, it is extremely difficult to model nonlinearities in this context.[7] As a result, we consider an approximating linear representation of the form:

$$y_t = a + \sum_{i=1}^{N} \beta_i x_{it} + u_t, \tag{2}$$

[6] The accompanying .xlsx file denotes the manual pre-selection with 1 and 0 in the column “Smaller.Selection”.

[7] However, in the next section we also include nowcasts based on random forests, neural networks and support vector regression which are able to capture non-linear relationships.

where u_t denotes a martingale difference process; depending on the estimation method this assumption can be relaxed to the standard assumptions in the classical linear regression framework.

3.1 Best Subset Selection

The first modelling approach is the best subset selection. Under this framework, the applied researcher aims to evaluate all possible models and choose the one which optimises a selected statistic; this statistic can be a measure of fit, such as R^2 or an information criterion, or can be an error-based statistics such as the Means Squared Error (MSE).

The main difficulty with this approach is that there are 2^N model combination to be evaluated. This means that if we have a small universe of indicators, say $N = 10$, we would have to evaluate 1,024 models which is easily doable with modern CPU power. This problem can be solved by adopting the "Best Forward Stepwise" (BFS) regressions.

In the BFS approach, we start with the null model:

$$y_t = a + u_t, \tag{3}$$

which includes only a constant. Then, we estimate the model by adding the first available variable, say x_{1t}:

$$y_t = a + \beta_1 x_{1t} + u_t, \tag{4}$$

and evaluate the chosen statistic of interest. We adopt the model which optimises the chosen statistic and then proceed with the next variable. In what follows we adopt Schwarz (1978) Bayesian Information Criterion (BIC).

It is important to highlight that this approach can provide a first variable selection and indicate the relative importance of indicators, however it might miss information hidden in the cross-section of variables which are not sequential to each other.

3.2 Penalised Regressions

Penalised regression is one of the most popular ways for sparse regression in the literature which, depending on the nature of the penalty, also allows for variable selection without having to consider the ordering of the variables. Various penalties have been suggested in order to effectively estimate the β_i parameters assigning zeros to the variables which should not be used in the regression (meaning that these are not part of the true model) and, consequently, in the nowcasting exercise. In what follows we denote $\beta_N = (\beta_1, ..., \beta_N)'$ and $x_N = (x_1, ..., x_N)'$.

3.2.1 Ridge Regression

Ridge Regression creates a linear regression model that is penalised with the L2-norm which is the sum of the squared coefficients. This has the effect of shrinking the coefficient values (and the complexity of the model) allowing some coefficients with minor contribution to the response to get close to zero (but not exactly equal to zero). The parameter estimators, $\widehat{\beta}^{Ridge}$, are then computed by solving the following optimisation problem:

$$\min_{\beta_N} \left\{ \sum_{t=1}^{T} \left(y_t - a - \beta_N' x_{t,N} \right)^2 + \lambda \sum_{i=1}^{N} \beta_i^2 \right\}, \tag{5}$$

for given values of a and λ. λ is the penalty parameter. OLS corresponds to the no penalty case, where $\widehat{\beta}^{Ridge} \to \widehat{\beta}^{OLS}$ as $\lambda \to 0$. Also, it can be easily seen that $\widehat{\beta}^{Ridge} \to 0$ as $\lambda \to \infty$. By centering the columns of x, the intercept becomes $\widehat{\alpha} = \overline{y}$. Therefore, we typically center y, x_N and do not include the intercept term.

The variance and bias of the ridge regression estimator can be shown to be:

$$Var\left(\widehat{\beta}^{Ridge}\right) = \sigma^2 W x_N' x_N W$$
$$Bias\left(\widehat{\beta}^{Ridge}\right) = -\lambda W \beta$$

where $W = \left(x_N' x_N + \lambda I\right)^{-1}$. It can be shown that the total variance ($\sum_j Var\left(\widehat{\beta}_j\right)$)

is a monotone decreasing sequence with respect to λ, while the total squared bias ($\sum_j Bias^2\left(\widehat{\beta}_j\right)$) is a monotone increasing sequence with respect to λ. The standard OLS assumptions are also required for Ridge regression.

Since Ridge regression does not set coefficients exactly equal to zero (unless $\lambda \to \infty$, in which case they are all zero), it cannot perform variable selection and, even though it might perform well in terms of prediction accuracy, it does not offer a clear interpretation of the resulting nowcasts.

3.2.2 LASSO Regression

Least Absolute Shrinkage and Selection Operator (LASSO) creates a regression model that is penalised with the L1-norm which is the sum of the absolute coefficients. Because of the nature of this constraint, it tends to produce some coefficients that are exactly equal to 0 and, hence, gives more interpretable models. Simulation studies suggest that the LASSO enjoys some of the favourable properties of both subset selection and ridge regression. As originally noted by Tibshirani (1996), the lasso regression is better suited for predictor selection compared to the Ridge regression because the former method performs model/predictors selection keeping those variables which are more suitable for forecasting. The optimisation problem now becomes:

$$\min_{\beta_N}\left\{\sum_{t=1}^{T}\left(y_t - a - \beta_N^{'} x_{t,N}\right)^2 + \lambda\sum_{i=1}^{N}|\beta_i|\right\}. \tag{6}$$

Although we cannot write the explicit formulas for the bias and variance of the LASSO estimator, the general trend is that the bias increases as λ increases and the variance decreases as λ increases.

Following Bühlmann and van de Geer (2011), we summarise the key properties and corresponding assumptions for the LASSO. Considering the true model in Equation (2), it is:

$$\frac{1}{T}\sum_{t=1}^{T}\left(x_{t,N}\left(\widehat{\beta}^{LASSO} - \beta\right)\right)^2 = O_P\left(\sum_{i=1}^{N}|\beta_i|\sqrt{\log\left(N\right)/T}\right), \tag{7}$$

where $O_P(\cdot)$ is with respect to $N \geq T \to \infty$. This implies that we achieve consistency of prediction if $\sum_{i=1}^{N} |\beta_i| \ll \sqrt{T/\log(N)}$.()

Faster convergence rate and estimation error bounds with respect to the L1- or L2-norm can be achieved using the so-called oracle optimality condition:

$$\frac{1}{T}\sum_{t=1}^{T}\left(x_{t,N}\left(\widehat{\beta}^{LASSO}-\beta\right)\right)^2 = O_P\left(s_0\phi^{-2}\log(N)/T\right),$$
$$\sum_{i=1}^{N}\left|\widehat{\beta}_i^{LASSO}-\beta_i\right|^q = O_P\left(s_0^{1/q}\phi^{-2}\sqrt{\log(N)/T}\right),\ q=\{1,2\}, \tag{8}$$

where s_0 equals the true number of non-zero regression coefficients and ϕ^2 is the compatibility constant or restricted eigenvalue which is a number depending on the compatibility between the design and the L1-norm of the regression coefficient. The above rate is optimal up to the $\log(N)$ factor and the restricted eigenvalue ϕ^2.

Additionally to the oracle optimality and assuming the beta-min condition:

$$\min_{i\in S_0^c}|\beta_i| \gg \phi^{-2}\sqrt{s_0\log(N)/T},$$

we obtain the screening variable property:

$$P\left[\widehat{S}\supseteq S\right]\to 1 \quad (N\geq T\to\infty), \tag{9}$$

where $\widehat{S}=\{i;\widehat{\beta}_i^{LASSO}\neq 0, i=1,..,N\}$ and $S=\{i;\beta_i\neq 0, i=1,..,N\}$. Consistent variable selection then means

$$P\left[\widehat{S}=S\right]\to 1 \quad (N\geq T\to\infty). \tag{10}$$

3.2.3 Adaptive LASSO

Zou (2006) introduces the adaptive LASSO (A-LASSO) estimator where the L1-norms in the penalty are re-weighted. He shows that, if a reasonable initial estimator is available, under appropriate conditions, the A-LASSO correctly selects covariates with nonzero coefficients with probability converging to one, and that the estimators of nonzero coefficients have the same asymptotic distribution they would have if the zero coefficients were known in advance.

The optimisation problem now becomes:

$$\min_{\beta_N}\left\{\sum_{t=1}^{T}\left(y_t-a-\beta_N^{'}x_{t,N}\right)^2+\lambda\sum_{i=1}^{N}\widehat{w}_i\left|\beta_i\right|\right\}, \tag{11}$$

where $\widehat{w}_i = 1/|\ \widehat{\beta}_{init,i}|^{\gamma}$, $\widehat{\beta}_{init}$ is an initial estimator and $\gamma > 0$. Usually, the initial estimator is the LASSO estimator with the constraint parameter tuned in the usual way with CV scheme as discussed earlier. Then, in the second stage CV is again used to select the λ parameter in Equation (11).

Following Haung et al. (2008) we consider the following conditions to hold for the variable selection and asymptotic normality of the A-LASSO in large samples.

1. The errors are iid.

2. The initial estimators $\widehat{\beta}_{init,i}$ are r_T-consistent for the estimation of certain η_{Ti}:

$$r_T\max_{i\leq N}\left|\widehat{\beta}_{init,i}-\eta_{T,i}\right|=O_P\left(1\right),\ r_T\rightarrow\infty$$

   where η_{Ti} are unknown constants depending on β_N and satisfy

$$\max_{i\notin J_{T1}}\left|\eta_{T,i}\right|\leq M_{T2},\ \left\{\sum_{i\in J_{T1}}\left(\frac{1}{\left|\eta_{Ti}\right|}+\frac{M_{T2}}{\left|\eta_{Ti}\right|^2}\right)^2\right\}^{1/2}\leq M_{T1}=o\left(r_T\right).$$

3. Adaptive irrepresentable condition. For $s_{T1}=\left(\left|\eta_{Ti}\right|^{-1}sgn\left(\beta_i\right),i\in J_{T1}\right)^{'}$ and some $\kappa<1$

$$\frac{1}{T}\left|x_i^{'}X_1\sum_{T11}^{-1}s_{T1}\right|\leq\frac{\kappa}{\left|\eta_{Ti}\right|},\forall i\notin J_{T1}.$$

4. The constants $\{k_T, m_T, \lambda_T, M_{T1}, M_{T2}, b_{T1}\}$ satisfy

$$(\log T)^{I\{d=1\}} \left\{ \frac{(\log k_T)^{1/d}}{T^{1/2} b_{T1}} + (\log m_T)^{1/d} \frac{T^{1/2}}{\lambda_T} \left(M_{T2} + \frac{1}{r_T} \right) \right\} + \frac{M_{T1}\lambda_T}{b_{T1} T} \to 0.$$

5. There exists a constant $\tau_1 > 0$ such that $\tau_{T1} \geq \tau_1$ for all T.

Following Haung et al. (2008), Condition 1 is standard for variable selection in linear regression. Condition 2 assumes that the initial $\widehat{\beta}_{init,i}$ actually estimates some proxy $\eta_{T,i}$ of β_i so that the weights are not too large for $\beta_{0i} \neq 0$ and not too small for $\beta_{0i} = 0$. The adaptive irrepresentable condition becomes the strong irrepresentable condition for the sign-consistency of the Lasso if the $|\eta_{T,i}|$ are identical for all $i \leq N$. Condition 4 restricts the numbers of covariates with zero and nonzero coefficients, the penalty parameter, and the smallest non-zero coefficient. Condition 5 assumes that the eigenvalues of Σ_{T11} are bounded away from zero, which is reasonable since the number of nonzero covariates is small in a sparse model. If the above conditions hold, then $P\left[\widehat{\beta}^{A-LASSO} = \beta\right] \to 1$.

3.2.4 Elastic Net

Elastic Net (EN) creates a regression model that is penalised with both the L1-norm and L2-norm. Introduced by Zou and Hastie (2005), the elastic net has the effect of effectively shrinking coefficients (as in ridge regression) and setting some coefficients to zero (as in LASSO). The optimisation problem now is:

$$\widehat{\beta}^{naiveEN} = \min_{\beta_N} \left\{ \sum_{t=1}^{T} \left(y_t - a - \beta_N^{'} x_{t,N} \right)^2 + \lambda_1 \sum_{i=1}^{N} |\beta_i| + \lambda_2 \sum_{i=1}^{N} \beta_i^2 \right\}. \tag{12}$$

The above is called the naive elastic net. A correction which leads to the elastic net is then:

$$\widehat{\beta}^{EN} = (1 + \lambda_2)\, \widehat{\beta}^{naiveEN}.$$

The correction factor $(1 + \lambda_2)$ is best motivated from the orthonormal design where $\frac{1}{T} x_N^{'} x_N = I$. The main advantage of the elastic net is its usefulness when the number of predictors is much bigger than the number of observations, which is usually the case in our big data context.

The reason for adding an additional squared L2-norm penalty is motivated by Zou and Hastie (2005) as follows. For strongly correlated covariates, the LASSO may select one but typically not both of them (and the non-selected variable can then be approximated as a linear function of the selected one). From the point of view of sparsity, this is what we would like to do. However, in terms of interpretation, we may want to have two even strongly correlated variables among the selected variables: this is motivated by the idea that we do not want to miss a "true" variable due to selection of a "non-true" which is highly correlated with the true one.

3.3 Factor Extraction via PCA

Another set of methods for modelling with a large panel of indicators involves the adoption of dimension reduction via techniques which do not require or impose any iid assumptions. In what follows we discuss the Principal Components Analysis (PCA) which has dominated the literature. This method is also frequently used for creating composite indicators (see Kapetanios and Papailias, 2021, for a recent discussion).

Factor methods have been at the forefront of developments in forecasting with large data sets and in fact started this literature with the influential work of Stock and Watson (2002a). The defining characteristic of most factor methods is that relatively few summaries of the large data sets are used in forecasting equations, which thereby become standard forecasting equations as they only involve a few explanatory variables.

The main assumption is that the co-movements across the indicator variables x_t, where $x_t = (x_{1t} \cdots x_{Nt})'$ is a vector of dimension $N \times 1$, can be captured by a $r \times 1$ vector of unobserved factors $F_t = (F_{1t} \cdots F_{rt})'$:

$$\tilde{x}_t = \Lambda' F_t + e_t \tag{13}$$

where $\tilde{x}_t$ may be equal to x_t or may involve other variables, such as lags or products of the elements of x_t, and Λ is an $r \times N$ matrix of parameters describing how the individual indicator variables relate to each of the r factors, which we denote with the terms 'loadings'. In Equation (13) e_t is a zero-mean $I(0)$ vector of errors that represent, for each indicator variable, the fraction of dynamics unexplained by F_t, the 'idiosyncratic components'. The number of factors is assumed to be finite. So, implicitly, in Equation (2) $\alpha' = \tilde{\alpha}'\Lambda\tilde{x}_t$, where $F_t = \Lambda\tilde{x}_t$, which means that a small, r, number of linear combinations of $\tilde{x}_t$ represent the factors and act as the predictors for y_t, the target variable. The main difference between different factor methods relates to how Λ and the factors are estimated.

The use of PCA for the estimation of factor models is, by far, the most popular factor extraction method. It has been popularised by Stock and Watson (2002a, 2002b), in the context of large data sets, although the idea had been well established in the traditional multivariate statistical literature. The method of principal components is simple. Estimates of Λ and the factors F_t are obtained by solving:

$$V(r) = \min_{\Lambda,F} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (\tilde{x}_{it} - \lambda_i' F_t)^2, \tag{14}$$

where λ_i is an $r \times 1$ vector of loadings that represent the N columns of $\Lambda = (\lambda_1 \cdots \lambda_N)$. One, non-unique, solution of Equation (14) can be found by taking the eigenvectors corresponding to the r largest eigenvalues of the second moment matrix $X'X$, which then are assumed to represent the rows in Λ, and the resulting estimate of Λ provides the forecaster with an estimate of the r factors $\hat{F}_t = \hat{\Lambda}\tilde{x}_t$. To identify the factors up to a rotation, the data are usually normalised to have zero mean and unit variance prior to the application of principal components; see Stock and Watson (2002a) and Bai (2003). We note that factor estimates obtained via PCA estimation are $\min(\sqrt{N}, T)$-consistent. Further, if $\sqrt{T}/N = o(1)$, using estimated factors rather than true factors in predictive regressions produces negligible estimation errors. PCA estimation of the factor structure is essentially a static exercise as no lags of x_t are considered.

3.4 Factor Extraction via PLS

Partial Least Squares (PLS) is a relatively new method for estimating regression equations, introduced in order to facilitate the estimation of multiple regressions when there is a large, but finite, amount of regressors.[8] The basic idea is similar to PCA in that factors or components, which are linear combinations of the original regression variables, are used instead of the original variables as regressors. PLS regression does not seem to have been explicitly considered for datasets with a very large number of series, i.e., when N is assumed in the limit to converge to infinity.

There is a variety of definitions for PLS and accompanying specific PLS algorithms that inevitably have much in common. A conceptually powerful way of defining PLS is to note that the PLS factors are those linear combinations of x_t, denoted by Υx_t, that give maximum covariance between y_t and Υx_t while being orthogonal to each other. Of course, in analogy to PCA factors, an identification assumption is needed to construct PLS factors in the usual form of normalisation.

A simple algorithm to construct the first k PLS factors is discussed among others, in detail, in Helland (1990). Assuming for simplicity that y_t has been demeaned and x_t have been normalized to have zero mean and unit variance, a simplified version of the algorithm is given below.

1. Set $u_t = y_t$ and $v_{i,t} = x_{i,t}$, $i = 1, ...N$. Set $j = 1$.

2. Determine the $N \times 1$ vector of indicator variable weights or loadings $w_j = (w_{1j} \cdots w_{Nj})'$ by computing individual covariances: $w_{ij} = Cov(u_t,\ v_{it})$, $i = 1, ..., N$. Construct the j-th PLS factor by taking the linear combination given by $w_j' v_t$ and denote this factor by $f_{j,t}$.

3. Regress u_t and $v_{i,t}$, $i = 1, ..., N$ on $f_{j,t}$. Denote the residuals of these regressions by $\tilde{u}_t$ and $\tilde{v}_{i,t}$ respectively.

[8] Herman Wold and co-workers introduced PLS regression between 1975 and 1982, see, e.g., Wold (1982). Since then it has received much attention in a variety of disciplines, especially in chemometrics, outside of economics.

4. If $j = k$ stop, else set $u_t = \tilde{u}_t$, $v_{i,t} = \tilde{v}_{i,t}$ $i = 1, .., N$ and $j = j + 1$ and go to step 2.

This algorithm makes clear that PLS is computationally tractable for very large data sets. Once PLS factors are constructed y_t can be modeled or forecast by regressing y_t on $f_{j,t}$, $j = 1, ..., k$. Helland (1990) provides a general description of the partial least squares (PLS) regression problem. Helland (1990) shows that the estimates of the coefficients α in the regression of y_t on x_t, as in Equation (2), obtained implicitly via PLS Algorithm and a regression of y_t on $f_{j,t}$ $j = 1, ..., k$, are mathematically equivalent to

$$\widehat{\beta}_{PLS} = V_k(V_k'X'XV_k)^{-1}V_k'X'y \tag{15}$$

with $V_{k_1} = (X'y \quad X'XX'y \quad \cdots \quad (X'X)^{k-1}X'y)$, $X = (x_1 \cdots x_T)'$ and $y = (y_1 \cdots y_T)'$. Thus, Equation (15) suggests that the PLS factors that result from the PLS Algorithm span the Krylov subspace generated by $X'X$ and $X'y$, resulting in valid approximations of the covariance between y_t and x_t.

A major difference between PCA and PLS is that, whereas in PCA regressions the factors are constructed taking into account only the values of the x_t variables, in PLS the relationship between y_t and x_t is considered as well in constructing the factors.

Recently, Kelly and Pruit (2015) and Groen and Kapetanios (2016) have extended and provided theoretical results on PLS showing that it can be also applied in the large N context, while Hepenstrick and Marcellino (2019) have introduced the mixed frequency version and provided empirical evidence in favour of its use for nowcasting with very large datasets.

3.5 Random Forests

In the above sections, we have reviewed methods based on the specification of a parametric model, typically a linear regression, which links the target variable y with a, possibly big, number of explanatory variables x. Regression trees are based on a partition of the space of the dependent variable y into M subsets R_m, with

y allocated to each subset according to a given rule and modelled as a different constant c_m in each subset. This is a powerful idea, since it can fit various functional relationship between y and a set of explanatory variables x, say $y = f(x)$, without imposing linearity or additivity, which are commonly assumed in standard linear regression models. Let

$$y = f(x) = \sum_{m=1}^{M} c_m \mathbf{1}\left(x \in R_m\right),$$

where $\mathbf{1}$ denotes the indicator variable taking value 1 if the condition is satisfied, 0 otherwise. Then, given a partition, minimising:

$$\|y - f(y)\|_2 = \sum_{i=1}^{N} \left(y_i - f(y_i)\right)^2, \tag{16}$$

with respect to c_m yields $\hat{c}_m = \overline{y}_m$, where $\overline{y}_m$ denotes the sample mean of y over each region R_m.

A much more difficult problem is to find the best partition in terms of minimum sum of squares. Even in the two dimensional case, i.e when $N = 2$ so that $X = [x_1, x_2]$, finding the best binary partition to minimise the sum of squares is not computationally feasible. Instead, greedy algorithms are commonly used. The idea is to do one split at a time. Consider a splitting variable j (where $j = 1, .., k$) and a splitpoint s such that a region $R_1(j, s)$ is defined as

$$R_1(j, s) = \{X|X_j \leq s\} \text{ and } R_2(j, s) = \{X|X_j > s\}$$

Then, the sum of squares) is minimised with respect to j and s. For each splitting variable, the split point s can be found and, hence, by scanning through all of the variables X_j, determination of the best pair (j, s) is feasible. Having found the best split, the data are partitioned into **two** resulting regions and the same splitting exercise is repeated on each of the two regions. Then this process is repeated on all of the resulting regions and so on. How many rounds of the algorithm are done determines how deep the resulting tree is. On one hand, shallow trees might fail to capture the structure of the data. On the other hand, however, deeper trees might overfit the data and, hence, do poorly in prediction.

There are ways to further improve the performance of regression trees. Bootstrap requires choosing with replacement a subsample and re-estimating the tree in order to get a sampling distribution of various statistics. Bagging is a general method that generates multiple versions of a predictor and uses these to get an aggregated predictor. Breiman (1996) offers an overview. In the context of regression trees, bagging averages across trees is estimated with different bootstrapped samples.

Random forests were introduced by Breiman (2001). The idea is exactly as bagging applied on regression trees: to grow a large collection of de-correlated trees (hence, the name forest) and then average them. This is achieved by bootstrapping a random sample at each node of every tree. In order to induce "decorrelation" of trees, when growing trees, before each split, select a subset of the input variables at random as candidates for splitting. This prevents the "strong" predictors imposing too much structure on the trunk of the tree.

3.6 Neural Networks

We continue with machine learning-based nonlinear methodologies and include two artificial neural network (ANN)-based approaches: (i) the Multilayer Perceptron (MLP) neural networks, and the (ii) Extreme learning machines (ELM) neural networks. Both MLP and ELM are feedforward artificial neural network with -at least-three layers of nodes: an input layer, a hidden layer (or multi-layers) and an output layer. Except for the input nodes, each node is a neuron that uses a nonlinear activation function. Both MLP and ELM are supervised machine learning approaches. In the following subsections, we provide some introduction to these approaches discussing their advantages and disadvantages. However, as it is beyond the scope of this task, we refer the reader to Ord et al. (2017) for more theoretical details on neural networks.

Both MPL and EML use the sigmoid functions in the neurons. Learning occurs in both methods by changing connection weights after each piece of data is processed, based on the amount of error in the output compared to the expected result. This is an example of supervised learning and is usually carried out through backpropagation. However, in the two neural networks incarnations we apply here we use the lasso approach instead of backpropagation.

3.6.1 Multilayer Perceptron

MLP is one of different kinds of ANNs. The term MLP can be used to loosely describe any feedforward ANN, sometimes strictly referring to networks composed of multiple layers of perceptrons (with threshold activation). MLPs with a single hidden layer are usually called "vanilla" neural networks; this is also the case in our empirical results later.

If a multilayer perceptron has a linear activation function in all neurons, that is, a linear function that maps the weighted inputs to the output of each neuron, then linear algebra shows that any number of layers can be reduced to a two-layer input-output model. In MLPs we can have some neurons to use a non-linear activation function.

The most commonly used form of ANNs for forecasting is the feedforward multilayer perceptron. The one-step ahead forecast $\hat{y}_{t+1}$ is computed using inputs that are lagged observations of the time series or other explanatory variables. I denotes the number of inputs p_i of the ANN. Its functional form can be described as:

$$\hat{y}_{t+1} = \alpha + \sum_{h=1}^{H} \beta_h g \left(\gamma_{0i} + \sum_{i=1}^{I} \gamma_{hi} p_i \right). \tag{17}$$

In the equation above, $\boldsymbol{w} = (\boldsymbol{\beta}, \boldsymbol{\gamma})$ are the network weights with $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_H]$ and $\gamma = [\gamma_{11}, \ldots, \gamma_{HI}]$ being the output and the hidden layers respectively. α and γ_{0i} are the biases of each neuron, which for each neuron act similarly to the intercept in a regression. H is the number of hidden nodes in the network and $g(\cdot)$ is a non-linear transfer function which is usually either the sigmoid logistic or the hyperbolic tangent function. ANNs can model interactions between inputs, if any. The outputs of the hidden nodes are connected to an output node that produces the forecast.

In the time series forecasting context, neural networks can be perceived as equivalent to nonlinear autoregressive models. Lags of the time series, potentially together with lagged observations of explanatory variables, are used as inputs to the network. During training, pairs of input vectors and targets are presented to the network. The network output is compared to the target and the resulting error is used to update the network weights. ANN training is a complex nonlinear optimisation problem and the network can often get trapped in local minima of the error surface. In order to avoid poor quality results, training should be initialised several times with different random starting weights and biases to explore the error surface more fully.

Note that the objective of training is not to identify the global optimum. This would result in the model over-fitting to the training sample and would then generalise poorly to unseen data, in particular given their powerful approximation capabilities. Furthermore, as new data becomes available, the prior global optimum may no longer be an optimum. In general, as the fitting sample changes, with the availability of new information, so do the final weights of the trained networks, even if the initial values of the network weights were kept constant. This sampling-induced uncertainty can again be countered by using ensembles of models, following the concept of bagging.

Advantages

1. MLPs are useful in research for their ability to solve problems stochastically, which often allows approximate solutions for extremely complex problems like fitness approximation.

2. MLPs are universal function approximators as shown by Cybenko's theorem, so they can be used to create mathematical models by regression analysis. As classification is a particular case of regression when the response variable is categorical, MLPs make good classifier algorithms.

3. MLPs were a popular machine learning solution in the 1980s, finding applications in diverse fields such as speech recognition, image recognition, and ma-

chine translation software, but thereafter faced strong competition from much simpler (and related) support vector machines.

Disadvantages

1. Computations are difficult and often time consuming.

2. The proper functioning of the model depends on the quality of the training.

3.6.2 Extreme Learning Machines

ELMs are feedforward neural networks for classification, regression, clustering, sparse approximation, compression and feature learning with a single layer or multiple layers of hidden nodes, where the parameters of hidden nodes need not be tuned. These hidden nodes can be randomly assigned and never updated, i.e. they are random projections but with nonlinear transforms, or can be inherited from their ancestors without being changed. In most cases, the output weights of hidden nodes are usually learned in a single step, which essentially amounts to learning a linear model.

Given a training set $\aleph = \{(\mathbf{x}_i, \mathbf{t}_i) \mid \mathbf{x}_i \in \ \mathbf{R}^n, \mathbf{t}_i \in \mathbf{R}^m, i = 1, \ldots, N\}$, activation function $g(x)$, and hidden node number $\tilde{N}$, we can describe the ELM generic algorithm as follows.

Step 1 : Randomly assign input weight $\mathbf{w}_i$ and bias b_i, $i = 1, \ldots, \tilde{N}$,

Step 2 : Calculate the hidden layer output matrix $\mathbf{H}$,

Step 3 : Calculate the output weight β defined as:

$$\beta = \mathbf{H}^{\dagger}\mathbf{T}, \tag{18}$$

where $\mathbf{T} = [\mathbf{t}_1, \ldots, \mathbf{t}_N]^{\mathrm{T}}$. In principle, this algorithm works for any infinitely differential activation function $g(x)$. Such activation functions include the sigmoidal functions as well as the radial basis, sine, cosine, exponential, and many non-regular functions. The upper bound of the required number of hidden nodes is the number of distinct training samples, that is $\tilde{N} \leqslant N$.

Advantages & Disadvantages. The black-box character of neural networks, in general, and ELMs in particular is one of the major concerns that repels engineers from application in unsafe automation tasks. This particular issue was approached by means of several different techniques. One approach is to reduce the dependence on the random input. Another approach focuses on the incorporation of continuous constraints into the learning process of ELMs which are derived from prior knowledge about the specific task. This is reasonable, because machine learning solutions have to guarantee a safe operation in many application domains. The mentioned studies revealed that the special form of ELMs, with its functional separation and the linear read-out weights, is particularly well suited for the efficient incorporation of continuous constraints in predefined regions of the input space.

3.7 Support Vector Regression

The final methodology we include in this report the Support Vector Regression (SVR). It can be argued that SVR is the adapted form of Support Vector Machines when the dependent variable is numerical rather than categorical. A major benefit of using SVR is that it is a non-parametric technique. Unlike the classical linear regression which depends on the Gauss-Markov theorem, the output model from SVR does not depend on distributions of the underlying dependent and independent variables. Instead the SVR technique depends on the kernel function which is employed to captured the underlying possibly non-linear relationships.

The success of the SVR is based on the fit of the kernel functions in the data. The most popular kernel functions include: (i) the linear, (ii) the polynomial, (iii) the sigmoid, and (iv) the radial basis function. It is important to highlight that different kernel functions result to different models and, therefore, different estimations and subsequent nowcasts. However, it is beyond the scope of the current report to provide an extensive discussion on the advantages and disadvantages of each kernel function. In what follows, we stick to the radial basis function.

The kernel allows the SVR to find a fit and then the data is mapped to the original space. In our setup, SVR works as simple non-linear model described as:

$$y_t = WK(x_t, x) + b, \tag{19}$$

where W is the product of the coefficients and the resulting support vectors and b is the negative intercept. We refer the reader to Chang and Lin (2021) Python implementation and references therein.

Following Vert et al. (2004), the RBF kernel on two samples $\mathbf{x}$ and $\mathbf{x}'$, represented as feature vectors in some input space, is defined as:

$$K\left(\mathbf{x}, \mathbf{x}'\right) = \exp\left(-\frac{\left\|\mathbf{x} - \mathbf{x}'\right\|^2}{2\sigma^2}\right) \tag{20}$$

where $\left\|\mathbf{x} - \mathbf{x}'\right\|^2$ may be recognized as the squared Euclidean distance between the two feature vectors. σ is a free parameter. An equivalent definition involves a parameter $\gamma = \frac{1}{2\sigma^2}$:

$$K\left(\mathbf{x}, \mathbf{x}'\right) = \exp\left(-\gamma \left\|\mathbf{x} - \mathbf{x}'\right\|^2\right). \tag{21}$$

Since the value of the RBF kernel decreases with distance and ranges between zero (in the limit) and one (when $\mathbf{x} = \mathbf{x}'$), it has a ready interpretation as a similarity measure. The feature space of the kernel has an infinite number of dimensions; for $\sigma = 1$, its expansion is:

$$
\begin{aligned}
\exp\left(-\frac{1}{2}\left\|\mathbf{x}-\mathbf{x}'\right\|^2\right) &= \exp\left(\frac{2}{2}\mathbf{x}^\top\mathbf{x}' - \frac{1}{2}\|\mathbf{x}\|^2 - \frac{1}{2}\left\|\mathbf{x}'\right\|^2\right) \\
&= \exp\left(\mathbf{x}^\top\mathbf{x}'\right)\exp\left(-\frac{1}{2}\|\mathbf{x}\|^2\right)\exp\left(-\frac{1}{2}\left\|\mathbf{x}'\right\|^2\right) \\
&= \sum_{j=0}^{\infty}\frac{\left(\mathbf{x}^\top\mathbf{x}'\right)^j}{j!}\exp\left(-\frac{1}{2}\|\mathbf{x}\|^2\right)\exp\left(-\frac{1}{2}\left\|\mathbf{x}'\right\|^2\right) \\
&= \sum_{j=0}^{\infty}\sum_{\sum n_i=j}\exp\left(-\frac{1}{2}\|\mathbf{x}\|^2\right)\frac{x_1^{n_1}\cdots x_k^{n_k}}{\sqrt{n_1!\cdots n_k!}}\exp\left(-\frac{1}{2}\left\|\mathbf{x}'\right\|^2\right)\frac{x_1'^{n_1}\cdots x_k'^{n_k}}{\sqrt{n_1!\cdots n_k!}}.
\end{aligned}
\tag{22}
$$

Because support vector machines and other models employing the kernel trick do not scale well to large numbers of training samples or large numbers of features in the input space, several approximations to the RBF kernel (and similar kernels) have been introduced. Typically, these take the form of a function z that maps a single vector to a vector of higher dimensionality, approximating the kernel:

$$
\langle z(\mathbf{x}), z\left(\mathbf{x}'\right)\rangle \approx \langle\varphi(\mathbf{x}), \varphi\left(\mathbf{x}'\right)\rangle = K\left(\mathbf{x}, \mathbf{x}'\right) \tag{23}
$$

where φ is the implicit mapping embedded in the RBF kernel.

One way to construct such a z is to randomly sample from the Fourier transformation of the kernel as noted in Rahimi et al. (2007). Another approach uses the Nyström method to approximate the eigendecomposition of the Gram matrix K, using only a random sample of the training set as given in Williams and Seeger (2001).

4 Nowcasting Setup

In this section, we describe the setup of the out-of-sample cross-validation nowcasting exercise. In what follows, x_t denotes the vector of all available predictors.

4.1 Algorithm

Our nowcasting algorithm is structured in the following way.

1. Set the out-of-sample nowcasting dates. We use the last four years as out-of-sample. In the case of UHP variables this includes the years 2013 to 2016. For most of the PSP variables, this includes the years 2016 to 2019.

2. For each given out-of-sample date, t, collect the available data for y_t and x_t and mimic the pattern of availability by imposing missing values according to the corresponding .xlsx file. It is important to notice here that we use three missing value approaches:

    - MV1: which assumes that nowcasting takes place 6 months after the reference period, i.e. $t + 0.5$ (in annual terms).
    - MV2: which assumes that nowcasting takes place 12 months after the reference period, i.e. $t + 1$ (in annual terms).
    - MV1: which assumes that nowcasting takes at some point after the reference period where x_t has full *availability*; in principle, this corresponds to $t + \delta$ with $\delta >> 1$.

3. As described in the Data Description section, once the pseudo-available data is ready, we perform the following operations at *each* nowcasting date in an effort to exactly mimic how the applied researcher would have done the estimation back then.

    - First, each variable is checked for seasonalities. If seasonalities are detected they are removed using the Cleveland et al. (1990) STL decomposition method.
    - Then, quarterly and monthly (including weekly-turned-to-monthly) variables are translated to annual using the annual averages (a simplified Bridge equations approach).

- Any missing values due to publication lags are exrapolated using the exponential average of the past 2 years.
- Then, variables are transformed to stationarity (the accompanying .xlsx file provides the transformation codes for each variable).
- Outliers are removed (if required).
- Colinear variables are removed; we define “colinear” variables as an pair of variables with an absolute correlation greater than 85%. Even though most of the state-of-the-art machine learning methodologies we employ are able to handle colinearities, removing them beforehand ensures a more efficient estimation.

4. Assuming that y_t has a missing value at the bottom, which we actually need to nowcast, estimate the linear regression model of the generic type:

$$y_{t-k} = \alpha + \beta x_{t-k} + \varepsilon_t, \tag{24}$$

obtain $\widehat{\alpha}$ and $\widehat{\beta}$ and produce the nowcast estimate as: $\widehat{y}_t = \widehat{\alpha} + \widehat{\beta} x_t$.

5. Repeat Steps 2 to 4 for all out-of-sample dates recursively (i.e. increasing the sample size).

It is important to notice that the above nowcasting takes place in the annual frequency using the stationary transformations for both y_t and x_t. However, at each step the log growth estimates are translated back to levels and the nowcast error is calculated in levels.

4.2 Evaluation

The purpose of this framework is to utilise nowcasting regressions and compare various models. To facilitate the nowcasting comparison we report standard nowcasting error statistics the Root Mean Squared Error (RMSE) evaluated for Model M across the out-of-sample nowcasts defined as:

$$RMSE_M = \sqrt{\frac{1}{T_{out}} \sum_{l=1}^{T_{out}} e_M^2}, \tag{25}$$

where e_M is the vector of T_{out} nowcast errors for Model M. We report the above statistic relative to a standard AR(1) benchmark, i.e. values smaller than 1 indicate predictive gains of the given model against the benchmark. This also allows the cross-comparison across models.

4.3 Models

In the out-of-sample nowcasting exercise, we consider the following models.

- Univariate time series benchmark models: **AR(1)** and **AR(P)** where $\widehat{P}$ is chosen minimising the Akaike (1973) Information Criterion (AIC).
- **BFS** selecting variables from the whole dataset (BFS-L) as well as the smaller manually selected variables (BFS-S).
- **Ridge** using variables from the whole dataset (Ridge-L) as well as the smaller manually selected variables (Ridge-S).
- **Lasso** regression using variables from the whole dataset (Lasso-L) as well as the smaller manually selected variables (Lasso-S).
- **EN** regression using a mixing parameter of $\alpha = 0.5$ for illustration, however many more choices could be considered. We use variables from the whole dataset (EN05-L) as well as the smaller manually selected variables (EN05-S).
- Adaptive Lasso regressions using the Ridge weights in the penalty; using variables from the whole dataset (AdRidge-L) as well as the smaller manually selected variables (AdRidge-S).

- Adaptive Lasso regressions using the Lasso weights in the penalty; using variables from the whole dataset (AdLasso-L) as well as the smaller manually selected variables (AdLasso-S).

- Adaptive Lasso regressions using the EN weights in the penalty; using variables from the whole dataset (AdEN05-L) as well as the smaller manually selected variables (AdEN05-L).

- Linear regression using **PCA(1)** factor from based in variables from the whole dataset (PCA1-L) as well as the smaller manually selected variables (PCA1-S).

- Linear regression using **PCA(2)** factor from based in variables from the whole dataset (PCA2-L) as well as the smaller manually selected variables (PCA2-S).

- Linear regression using **PCA(3)** factor from based in variables from the whole dataset (PCA3-L) as well as the smaller manually selected variables (PCA3-S).

- Linear regression using **PLS(1)** factor from based in variables from the whole dataset (PLS1-L) as well as the smaller manually selected variables (PLS1-S).

- Linear regression using **PLS(2)** factor from based in variables from the whole dataset (PLS2-L) as well as the smaller manually selected variables (PLS2-S).

- Linear regression using **PLS(3)** factor from based in variables from the whole dataset (PLS3-L) as well as the smaller manually selected variables (PLS3-S).

- **Random Forests** using the Lasso weights in the penalty; using variables from the whole dataset (RF-L) as well as the smaller manually selected variables (RF-S).

- **MLP** using the Lasso weights in the penalty; using variables from the whole dataset (MPL-L) as well as the smaller manually selected variables (MPL-S).

- **ELM** using the Lasso weights in the penalty; using variables from the whole dataset (ELM-L) as well as the smaller manually selected variables (ELM-S).

- **SVR** using the Lasso weights in the penalty; using variables from the whole dataset (SVR-L) as well as the smaller manually selected variables (SVR-S).
- Average of the AR models.
- Average of the variable selection models: BFS, Ridge, Lasso, Elastic Net and their adaptive versions for the whole dataset (AVG-BFS-PEN-L) as well as the smaller manually selected variables (AVG-BFS-PEN-S).
- Average of the PCA models: PCA(1), PCA(2) and PCA(3) using the whole dataset (AVG-PCA-L) as well as the smaller manually selected variables (AVG-PCA-L).
- Average of the PLS models: PLS(1), PLS(2) and PLS(3) using the whole dataset (AVG-PLS-L) as well as the smaller manually selected variables (AVG-PLS-L).
- Average of the nonlinear machine learning based methodologies models: Random Forests, MLP, EML, SVR using the whole dataset (AVG-NL-L) as well as the smaller manually selected variables (AVG-NL-L).
- Average of all models excluding the autoregressive benchmarks using the whole dataset (AVG-ALL-L) as well as the smaller manually selected variables (AVG-ALL-L).
- Average of all models including the autoregressive benchmarks using the whole dataset (AVG-ALL-wAR-L) as well as the smaller manually selected variables (AVG-ALL-wAR-L).

PCA and PLS factor models could consider up to about 6 factors in the linear regression, however we limit it to the first three to allow enough degrees of freedom; it is important to remember that these regressions are based on a sample of 8 to 11 observations only.

It is important to highlight that our main purpose is not to identify the best methodology or model suitable for nowcasting. Instead, the suggested framework utilises various models to provide robust evidence on the predictive content of the specific set of indicators considered and conclude whether nowcasting in this context provides encouraging results.

5 Empirical Results

This section is concerned with a brief discussion of the empirical output of the nowcasting exercise evaluated in the past four years for each case. We distinguish three missing value patterns for the dataset of potential predictors considering that nowcasting takes place at $t+0.5$ year after the reference date, $t+1$ year after the reference date and $t+\delta$ where $\delta > 1$ year; the latter case corresponds to a full data availability for the set of predictors. In all cases we consider a one missing value for the target variable.

The cross-validation results are presented in Tables 4 to 17. This collection of tables also inclues results with some real out-of-sample estimates which correspond to recent years where official data is yet to be released (i.e. 2017 to 2022 for UHP and 2020 to 2022 for PSP). The Appendix provides some illustrative examples for a selection of models using the MV3 cross-validation (overlapping period across all lines) and the corresponding real out-of-sample estimates.

5.1 UHP

5.1.1 Nowcasting at $t + 0.5$

Table 4 presents the cross-validation results of the various models based on MV1 missing value pattern. The table presents all models using the full set of predictors (L) as well as the smaller manually selected set (S). Also, the table presents the results for the market-equivalent outputs (O1 to O8 variables) as well as the full market-equivalent intermediate consumptions (I1 variables to I7 variables) along with there simple aggregates (O-AGG and I-AGG).

A first quick look at the table reveals that in most of the cases apart from O8 and I4, the top model always outperform the univariate AR benchmarks. The model with the smallest nowcast error is AdEN05-L, EN05-S, PCA(3)-L, MLP-L, AdEN05-S, BFS-S, AVG-BFS-PEN-L, AR(P) and BFS-S for the O1 to O8 and O-AGG respectively. Similarly, the model with the smallest error for the I variables is AdLasso-S, AVG-BFS-PEN-S, PCA(1)-S, AR(P), BFS-L, BFS-L, PCA(2)-S and PCA(1)-S for I1 to I7 and I-AGG respectively.

It is also observed, as expected, heterogeneity for the best performing models across target variables. This is not surprising as each target has its own behaviour which is likely to be captured by different models. Of course, a larger number in the cross-validation periods could increase the robustness of these results. Overall, we see that across all O and I cases, the best performing model is a variable selection model (BFS, Sparse Regressions) in 58.8% of the times (i.e. in 10 out of the 17 cases), a factor model (PCA, PLS) is the best in 23.5% of the times (4 out of the 17 cases), a nonlinear machine learning model (RF, MLP, EML, SVR) is the best in 5.9% of the times (1 out of the 17 cases) and a univariate AR model is the best in 11.8% of the times (2 out of the 18 cases). Furthermore, the best model is based on the full set of predictors (L) in 35.3% of the times (i.e. in 6 out of the 17 cases) and the best model is based on the small dataset (S) in 52.9% of the times (9 out of the 17 cases).

Finally, comparing the estimates by targeting directly the total aggregate (top only) and aggregating the estimates of the individual components (bottom up) we find that this works only in 18.5% of the models (9 out of 49 models provide improved estimates) for the O target variables. However, it seems to work better for the I targets as 53.1% of the models (26 out of 49) return improved estimates.

5.1.2 Nowcasting at $t + 1$

Table 5 presents the cross-validation results of the various models based on MV2 missing value pattern. As before, the table presents all models using the full set of predictors (L) as well as the smaller manually selected set (S). Also, the table presents the results for the market-equivalent outputs (O1 to O8 variables) as well as the full market-equivalent intermediate consumptions (I1 variables to I7 variables) along with there simple aggregates (O-AGG and I-AGG).

A first quick look at the table reveals that in most of the cases apart from O5, the top model always outperform the univariate AR benchmarks. The model with the smallest nowcast error is PCA(2)-L, BFS-L, PCA(3)-L, MLP-L, AR(1), BFS-S, AVG-BFS-PEN-L, AdRidge-S and BFS-S for the O1 to O8 and O-AGG respectively. Similarly, the model with the smallest error for the I variables is Lasso-L, AVG-BFS-PEN-S, BFS-L, Lasso-L, BFS-L, BFS-L, PCA(2)-S and PCA(1)-S for I1 to I7 and I-AGG respectively.

It is also observed, as expected, heterogeneity for the best performing models across target variables. Overall, we see that across all O and I cases, the best performing model is a variable selection model (BFS, Sparse Regressions) in 70.6% of the cases, a factor model (PCA, PLS) is the best in 23.5% of the cases, a nonlinear machine learning model (RF, MLP, EML, SVR) is the best in 5.9% of the cases and a univariate AR model is the best in 5.9% of the cases. Contrary to the previous case, the best model is based on the full set of predictors (L) in 58.8% of the cases and the best model is based on the small dataset (S) in 35.5% of the cases.

Finally, comparing the estimates by targeting directly the total aggregate (top only) and aggregating the estimates of the individual components (bottom up) we find that this works in 22.4% of the models for the O target variables. However, it seems to work better for the I targets as 46.9% of the models return improved estimates.

5.1.3 Nowcasting at Future Dates

Table 6 presents the cross-validation results of the various models based on MV3 missing value pattern which assumes that nowcasting takes place at a later date, $t+\delta$ where $\delta > 1$ year, when all data is available in the set of the potential predictors. As before, the table presents all models using the full set of predictors (L) as well as the smaller manually selected set (S). Also, the table presents the results for the market-equivalent outputs (O1 to O8 variables) as well as the full market-equivalent intermediate consumptions (I1 variables to I7 variables) along with there simple aggregates (O-AGG and I-AGG).

A first quick look at the table reveals that in most of the cases apart from O5 and O8, the top model always outperform the univariate AR benchmarks. The model with the smallest nowcast error is PCA(1)-S, AdEN05-L, AdRidge-S, MLP-S, AR(1), PCA(3)-S, EML-L, AR(P) and BFS-L for the O1 to O8 and O-AGG respectively. Similarly, the model with the smallest error for the I variables is AdLasso-L, AVG-PCA-L, AdRidge-S, Lasso-L, AdLasso-S, AdLasso-S, PCA(3)-S and AdLasso-S for I1 to I7 and I-AGG respectively.

It is also observed, as expected, heterogeneity for the best performing models across target variables. Overall, we see that across all O and I cases, the best performing model is a variable selection model (BFS, Sparse Regressions) in 52.9% of the cases, a factor model (PCA, PLS) is the best in 29.4% of the cases, a nonlinear machine learning model (RF, MLP, EML, SVR) is the best in 11.8% of the cases and a univariate AR model is the best in 11.8% of the cases. Furthermore, the best model is based on the full set of predictors (L) in 35.3% of the cases and the best model is based on the small dataset (S) in 52.9% of the cases.

Finally, comparing the estimates by targeting directly the total aggregate (top only) and aggregating the estimates of the individual components (bottom up) we find that this works only in 18.4% of the models for the O target variables. However, it seems to work better for the I targets as 46.9% of the models (26 out of 49) return improved estimates.

The hereby provided methodological tool returns the variables selected by various variable selection models and the loadings of various factor extraction models at each run in the cross-validation periods. The interested reader can inspect the detailed output to investigate the variables which seem to contribute the most in the estimation process. Indicatively, we choose to present the results for O1 (the market-equivalent output of household housing services provided) and the corresponding aggregate, O-AGG.

Table 7 presents the loadings of the variables considered by the PCA(1)-S regression model for O1 across the four years of the evaluation period. To investigate the contribution of each variable to the PCA(1) factor, one can rank the variables by their absolute loadings. In that way, we see that gbbank02092 (England & Wales, Bankruptcies, Total New Company Insolvencies, SA) is included in the Top 5 variables in 3 out of the 4 evaluation periods. It is followed by gbdemo0232 (Households, Three People) and lse_all_mcap (Equity Statistics, All Sections, London Stock Exchange, Market Cap, GBP) which are included in the variables with the Top 5 absolute loadings in 2 out of the 4 evaluation periods. This selection of variables does make sense as the household services are seemed to be tracked by the number of households with three people and the market cap in LSE (possibly as a measure of the overall growth).

In a similar manner we can investigate the variables selected by BFS-L for the O-AGG case. As we can see in Table 8, BFS includes in the Top 5 variables ranked by their absolute beta coefficient estimates the gbdemo0236 (Households, Total) and gb_rl_est_esg (World Bank, Rule of Law Index) in 3 out of the 4 evaluation periods. Then, ndgain_vul_1143 (Climate Change, Vulnerability Scores, Infrastructure, Index), gbbopa00482 (Capital Account, Total, Credit, SA) and gbpric00181 (Consumer Price Index, Health, Index) are included in the Top 5 variables in 2 out of 4 evaluation periods. Again, this selection of variables makes reasonable sense as the total aggregate of the UHP seems to be tracked by the number of households, the rule of law, the expensiveness (as proxied by the level of prices) and the total credit (which is argued to be an important indicator for economic crises[9]).

Using the above results, one could perform a real out-of-sample estimation and nowcast the values from 2017 to 2022 for the UHP O and I targets. Table 9 provides the estimates for the best model in the cross-validation. The figures in the Appendix provide illustrative examples of estimates and cross-validation results for some models for the interested reader.

[9]See Frankel and Saravelos (2010) and Babecký et al. (2013), among many others, for more details.

5.2 PSP

In this part, we turn our attention to the PSP target variables and their components. Due to data availability issues, we focus on Education (E) using 10 out of the 11 components (i.e. all apart from "Raw disadvantaged attainment gap index referenced to 100") and the Public Order and Safety (POS) three components (excluding magistrates index and crown index).

As, in this case, the final aggregate is expressed as a quality adjusted index -and not a direct aggregate, the comparison between top only (nowcasting the quality adjusted index directly) and the bottom up approach (aggregating the individual estimates of each components) involves the following two-step scaling procedure.

- In the first step, we use all the data for the individual components, denoted as $y_{j\tau}$ for the $j-th$ component for j =1, ..., K, as well as the final quality adjusted index, Y_τ, which is observed during τ =1, 2, ..., $t-1$, where t denotes the initial out-of-sample date in the evaluation period.

- We aggregate the individual components and create the aggregate variable defined as $y_\tau = \sum_{j=1}^{K} y_{j\tau}$. It is obvious that y_τ and Y_τ have scaling difference as y_τ is expected to taken large values (since it is the sum of all components) whereas Y_τ has 100 as a reference value.

- We estimate the following regression to level these differences: $Y_\tau = \alpha + \beta y_\tau + \varepsilon_\tau$ where $\varepsilon_\tau \sim WN(0, \sigma_\varepsilon^2)$ with $\sigma_\varepsilon^2 < \infty$, and obtain $\widehat{\alpha}$ and $\widehat{\beta}$.

- We aggregate the nowcast estimates for all j =1, ..., K individual components, $\widehat{y}_{jt}$, and adjust them using the previously obtained coefficient estimates: $\widehat{Y}_t = \widehat{\alpha} + \widehat{\beta} \sum_{j=1}^{K} \widehat{y}_{jt}$ where t denotes the out-of-sample evaluation periods. $\widehat{Y}_t$, which is the bottom up estimate, can then be compared to the nowcast estimates we obtain target directly the final quality adjusted index.

It is important to highlight that the above two-step procedure uses estimates in the *observed* sample a researcher has available prior to the nowcasting and, hence, does not consider any future values and mimics the behaviour of the researcher in real-time.

As the targets, and specifically the quality adjusted indices, do not have specific publication dates, it is often difficult to identify the number of missing values due to the publication delays. In an effort to mimic real conditions as close as possible, we use 2 missing values for the targets when the nowcasting takes place at time $t + 0.5$ and 1 missing value when the nowcasting takes pace at $t + 1$ and future dates.

5.2.1 Nowcasting at $t + 0.5$

Table 10 presents the cross-validation results of the various models based on MV1 missing value pattern. The table presents all models using the full set of predictors (L) as well as the smaller manually selected set (S). Also, the table presents the results for the education variables (E1 to E10 variables) as well as the public order and safety variables (POS1 variables to POS3 variables) along with their corresponding quality adjusted indices (E-AGG and POS-AGG).

A first quick look at the table reveals that in most of the cases apart from E6 and E9, the top model always outperform the univariate AR benchmarks. The model with the smallest nowcast error is PCA(3)-S, PCA(3)-L, AdRidge-S, PCA(3)-L, BFS-L, AR(1), BFS-S, MLP-S, AR(P), AVG-NL-S and MLP-S for the E1 to E10 and E-AGG respectively. Similarly, the model with the smallest error for the POS variables is BFS-S, PCA(3)-S, MLP-S and MLP-S for POS1 to POS3 and POS-AGG respectively.

It is also observed, as expected, heterogeneity for the best performing models across target variables. This is not surprising as each target has its own behaviour which is likely to be captured by different models. Of course, a larger number in the cross-validation periods could increase the robustness of these results. Overall, we see that across all E and POS cases, the best performing model is a variable selection model (BFS, Sparse Regressions) in 26.7% of the times (i.e. in 4 out of the 15 cases), a factor model (PCA, PLS) is the best in 26.7% of the times (4 out of the 15 cases), a nonlinear machine learning model (RF, MLP, EML, SVR) is the best in 33.3% of the times (5 out of the 15 cases) and a univariate AR model is the best in 13.3% of the times (2 out of the 15 cases). Furthermore, the best model is based on the full set of predictors (L) in 20% of the times (i.e. in 3out of the 15cases) and the best model is based on the small dataset (S) in 66.7% of the times (10 out of the 15 cases).

Finally, comparing the estimates by targeting directly the total aggregate (top only) and aggregating the estimates of the individual components (bottom up) we find that this works only in 2% of the models (1 out of 49 models provide improved estimates) for the E target variables and does not work at all for the POS variables.

5.2.2 Nowcasting at $t + 1$

Table 11 presents the cross-validation results of the various models based on MV2 missing value pattern in a similar format as that of Table 10. Looking at the table, we see that in most of the cases apart from E10 and E-AGG, the top model always outperform the univariate AR benchmarks. The model with the smallest nowcast error is AVG-BFS-PEN-S, BFS-S, PCA(2)-S, AdRidge-L, BFS-L, MLP-L, BFS-S, PCA(1)-L, BFS-S, AR(P) and AVG-AR for the E1 to E10 and E-AGG respectively. Similarly, the model with the smallest error for the POS variables is AdLasso-L, PCA(3)-S, AdEN05-S and EML-S for POS1 to POS3 and POS-AGG respectively.

Overall, we again observe heterogeneity in the top performing models across all cases. in particular, the best performing model is a variable selection model (BFS, Sparse Regressions) in 53.3% of the cases, a factor model (PCA, PLS) is the best in 20% of the cases, a nonlinear machine learning model (RF, MLP, EML, SVR) is the best in 13.3% of the cases and a univariate AR model is the best in 13.3% of the cases. The best model is based on the full set of predictors (L) in 33.3% of the cases and the best model is based on the small dataset (S) is also found to be 33.3% of the cases.

Finally, as in the previous case, comparing the estimates by targeting directly the total aggregate (top only) and aggregating the estimates of the individual components (bottom up) we find that this works only in 2% of the models (1 out of 49 models provide improved estimates) for the E target variables and does not work at all for the POS variables.

5.2.3 Nowcasting at Future Dates

Table 12 presents the cross-validation results of the various models based on MV3 missing value pattern which assumes that nowcasting takes place at a later date, $t+\delta$ where $\delta > 1$ year, when all data is available in the set of the potential predictors.

Looking at the table we see that in most of the cases apart from E-AGG, the top model always outperform the univariate AR benchmarks. The model with the smallest nowcast error is EN05-S, BFS-L, PCA(2)-L, PCA(1)-S, AdLasso-L, MLP-S, AdEN05-L, AVG-BFS-PEN-S, BFS-S, MLP-L and AVG-AR for the E1 to E10 and E-AGG respectively. Similarly, the model with the smallest error for the POS variables is BFS-S, PCA(3)-S, AdLasso-L and PCA(2)-L for POS1 to POS3 and POS-AGG respectively.

Overall, we see that across all E and POS cases, the best performing model is a variable selection model (BFS, Sparse Regressions) in 53.3% of the cases, a factor model (PCA, PLS) is the best in 26.7% of the cases, a nonlinear machine learning model (RF, MLP, EML, SVR) is the best in 13.3% of the cases and a univariate AR model is the best in 6.7% of the cases. Furthermore, the best model is based on the full set of predictors (L) in 46.7% of the cases and the best model is based on the small dataset (S) in 53.3% of the cases. Also, the bottom-up nowcasting approach does not seem to work at all in this case.

The hereby provided methodological tool returns the variables selected by various variable selection models and the loadings of various factor extraction models at each run in the cross-validation periods. The interested reader can inspect the detailed output to investigate the variables which seem to contribute the most in the estimation process. Indicatively, we choose to present the results for O1 (the market-equivalent output of household housing services provided) and the corresponding aggregate, O-AGG.

Table 13 presents the loadings of the variables considered by the EN05-S regression model for E1 across the four years of the evaluation period. To investigate the impact of each variable, we rank them according the the magnitude of their absolute beta values. In that way, we see that gbdemo1180 (England & Wales, Deaths, Male), gb_swiid_rel_red (Wealth & Poverty, Relative Redistribution, Index) and uspoli0211 (Heritage Foundation Economic Freedom Index) are included in the Top 5 variables in 4 out of the 4 evaluation periods. These are followed by gbfofi0103 (International Reserves, Gold) and gbmost0100 (Monetary Statistics, M4) which are included in the Top 5 variables in 3 out of the 4 evaluation periods.

In a similar manner we can investigate the variables selected by BFS-S for the E9 case. As we can see in Table 14, BFS includes in the Top 5 variables the lse_all_mcap (London Stock Exchange, Market Cap, GBP) as it is selected in 4 out of 4 evaluation rounds. This is followed by gb_pv_est_esg (World Bank, Political Stability & Absence of Violence/Terrorism: Estimate, Index) and ukepunnewsindex (Economic Policy Uncertainty, News Based Index, 11 Newspapers) which are included in the Top 5 in 3 out of 4 times.

Finally, Table 15 presents the selected variables by BFS-S for POS1 target. We see that bp_gbenergyconsej (BP Statistical Review of World Energy, Primary Energy, Consumption), gbdemo0233 (Households, Four People), gbbank6836 (Payment System, CHAPS) and gbpric00211 (Consumer Price Index, Recreation & Culture, Index) are in the Top 5 models in 2 out of the 4 evaluation periods.

Overall, we can see that the models are towards a reasonable direction to identify variables which are can be linked to the corresponding targets qualitative as well as quantitatively. However, the (very) limited number in the evaluation rounds -perhaps- does not allow the models to “train” enough.

Using the above results, one could perform a real out-of-sample estimation and nowcast the values from 2017 to 2022 for the PSP E and POS targets. Table 16 provides the estimates for the best model in the cross-validation. The figures in the Appendix provide illustrative examples of estimates and cross-validation results for some models for the interested reader.

5.2.4 Nowcasting the Quality Adjusted Indices

In this final part of our exercise, we investigate the nowcasting performance directly using the four final quality adjusted indices for education, adult social care (ASC), children's social care (CSC) and public order and safety.

Table 17 presents the cross-validation results for the four indices across the three missing value patterns. AR(1) and AVG-AR provide better estimates for CSC with MV1, E with MV2 and E with MV3. In all other cases, they are being outperformed by another model.

6 Concluding Remarks

This research provides a generic methodological framework to nowcast variables which are important aspects of people's lives but are largely missing from regular economic statistics such as the GDP. We employ a large panel of potential predictors and models based on state-of-the-art machine learning approaches considering models from the following methodologies: best subset selection, sparse regressions, factor extraction, random forests, neural networks and support vector regression; see Kapetanios and Papailias (2021, 2022a, 2022b). To the best of our knowledge, this is the first paper that attempts to nowcast this type of economic variables. The empirical results are encouraging indicating that there exist models which can outperform the standard benchmarks.

In particular, the limited nowcasting cross-validation exercise in UHP targets reveals that variable selection models (BFS, Sparse Regressions) provide better estimates in most of the cases. The full sample of predictors is better for nowcasting at about $t + 1$ and that the bottom-up approach (i.e. estimating each individual component and aggregating) works in about half of the models only for the I targets. It also makes economic sense to obtain poor performance results for variables such as O5/I5 (laundry) or O8 (voluntary services) given that the majority of the variables in the set of potential predictors do not seem to be related with such activities.[10]

[10]No relevant variables have been identified for these series, however -in case such variables exist-

Moving to the PSP variables, we see that for MV1 (where the target has two missing values) nonlinear machine learning models seem to provide smaller error. As we move to MV2 (where the target has one missing value), we again see that variable selection models (BFS, Sparse Regressions) provide better estimates in most of the cases. Furthermore, the bottom-up approach does not seem to work well for these variables.

Overall, we believe that this methodological toolkit provides adequate results analysis to be successfully used by the applied researcher.

the provided methodological tool can handle any dataset in a similar format.

7 References

1. Akaike, H., 1974. A New Look at the Statistical Model Identification. IEEE Transactions on Automatic Control, 19(6), 716–723.

2. Babecký, J., Havránek, T., Matějȯ, J., Rusnák, M., Šmídková, K., Vašíček, B., 2013. Leading Indicators of Crisis Incidence: Evidence from Developed Countries. Journal of International Money and Finance, 35, 1-19.

3. Bai, J., 2003. Inferential Theory for Factor Models of Large Dimension. Econometrica, 71, 135-173.

4. Boivin, J., Ng, S., 2006. Are More Data Always Better for Factor Analysis? Journal of Econometrics, 132(1), 169-194.

5. Breiman, L., 1996. Bagging Predictors. Machine Learning, 24(2), 123–140.

6. Breiman, L., 2001. Random Forests. Machine Learning, 45(1), 5-32.

7. Bühlmann, P., van de Geer, S., 2011. Statistics for High-Dimensional Data: Methods, Theory and Applications. Springer.

8. Caggiano, G., Kapetanios, G., Labhard, V., 2011. Are More Data Always Better for Factor Analysis? Results for the Euro Area, the Six Largest Euro Area countries and the UK. Journal of Forecasting, 30(8), 736-752.

9. Camba-Mendez, G., Kapetanios, G., Papailias, F., Weale, M.R., 2015. An Automatic Leading Indicator, Variable Reduction and Variable Selection Methods using Small and Large datasets: Forecasting the Industrial Production Growth for Euro Area Economies. ECB Working Paper Series, No 1773/April 2015.

10. Camba-Mendez, G., Kapetanios, G., Smith, R.J., Weale, M.R., 2001. An Automatic Leading Indicator of Economic Activity: Forecasting GDP Growth for European Countries. Econometrics Journal, Royal Economic Society, 4(1), 1-37.

11. Chang, C.-C., Lin, C.-J., 2021. LIBSVM – A Library for Support Vector Machines. Version 3.25, released on April 14, 2021.

12. Cleveland, R.B., Cleveland, W.S., McRae, J.E., Terpenning, I.J., 1990. STL: A Seasonal-Trend Decomposition Procedure Based on Loess. Journal of Official Statistics, 6(1), 3–33.

13. Frankel, J.A., Saravelos, G., 2010. Are Leading Indicators of Financial Crises Useful for Assessing Country Vulnerability? Evidence from the 2008-09 Global Crisis. NBER Working Paper No. 16047, June 2010.

14. Groen, J.J.J., Kapetanios, G., 2016. Revisiting Useful Approaches to Data-Rich Macroeconomic Forecasting. Computational Statistics & Data Analysis, 100, 221-239.

15. Haung, J., Ma, S., Zhang, C.H., 2008. Adaptive LASSO for Sparse High Dimensional Regression Models. Statistica Sinica, 18, 1603-1618.

16. Helland, I.S., 1990). PLS Regression and Statistical Models. Scandivian Journal of Statistics, 17, 97-114.

17. Hepenstrick, C., Marcellino, M., 2019. Forecasting Gross Domestic Product Growth with Large Unbalanced Data Sets: The Mixed Frequency Three-Pass Regression Filter, Journal of the Royal Statistical Society: Series A, 182, 69-99.

18. Kapetanios G., Papailias F., 2021. UK Economic Conditions during the Pandemic: Assessing the Economy using ONS Faster Indicators. ESCoE Working Paper, DP 2021-10. Available here.

19. Kapetanios G., Papailias F., 2022a. Real Time Indicators During the COVID-19 Pandemic Individual Predictors & Selection. ESCoE Technical Report, TR-15. Available here.

20. Kapetanios G., Papailias F., 2022b. An Evaluation Framework for Targeted Indicators: Aggregates vs. Disaggregates. ESCoE Technical Report TR-17. Available here.

21. Kelly, B., Pruitt, S., 2015. The Three-Pass Regression Filter: A New Approach to Forecasting using Many Predictors. Journal of Econometrics, 186(2), 294-316.

22. Ord, K., Fildes, R., Kourentzes, N. 2017. Principles of Business Forecasting, 2e. Wessex Press Publishing Co., Chapter 10.

23. Rahimi, A., Recht, B., 2007. Random Features for Large-Scale Kernel Machines. NIPS '07: Proceedings of the 20th International Conference on Neural Information Processing Systems. December, 2007, 1177–1184.

24. Schwarz, G.E., 1978. Estimating the Dimension of a Model. Annals of Statistics, 6(2), 461–464.

25. Stock, J.H., Watson, M.W., 2002a. Forecasting using Principal Components from a Large Number of Predictors. Journal of the American Statistical Association, 97, 147-162.

26. Stock, J.H., Watson, M.W., 2002b. Macroeconomic Forecasting using Diffusion Indexes. Journal of Business and Economic Statistics, 20, 147-162.

27. Tibshirani, R., 1996. Regression Shrinkage and Selection via the Lasso. Journal of the Royal Statistical Society: Series B, 58(1), 267-288.

28. Vert, J.P., Tsuda, K., Scholkopf, B., 2004. Kernel Methods in Computational Biology. MIT Press.

29. Williams, C.K.I., Seeger, M., 2001. Using the Nyström Method to Speed Up Kernel Machines. Advances in Neural Information Processing Systems 13 (NIPS 2000), MIT Press, 682-688.

30. Zou, H., 2006. The Adaptive Lasso and Its Oracle Properties. Journal of the American Statistical Association, 101(746), 1418-1429.

31. Zou, H., Hastie, T., 2005. Regularization and Variable Selection via the Elastic Net. Journal of the Royal Statistical Society: Series B, 67(2), 301-320.

Tables

Code	Category	Description
O1	Household housing services	Market-equivalent output of household housing services provided
O2	Transport	Market-equivalent output of unpaid transport services provided
O3	Nutrition	Market-equivalent output of unpaid nutrition services provided
O4	Clothing	Market-equivalent output of unpaid clothing services provided
O5	Laundry	Market-equivalent output of unpaid laundry services provided
O6	Childcare	Market-equivalent output of unpaid childcare services provided
O7	Adult Care	Market-equivalent output of unpaid adultcare services provided
O8	Voluntary activity	Market-equivalent output of unpaid volunteering services provided
I1	Household housing services	Full Market-equivalent intermediate consumption of household housing services provided
I2	Transport	Full Market-equivalent intermediate consumption of unpaid transport services provided
I3	Nutrition	Full Market-equivalent intermediate consumption of unpaid nutrition services provided
I4	Clothing	Full Market-equivalent intermediate consumption of unpaid clothing services provided
I5	Laundry	Full Market-equivalent intermediate consumption of unpaid laundry services provided
I6	Childcare	Full Market-equivalent intermediate consumption of unpaid childcare services provided
I7	Adult Care	Full Market-equivalent intermediate consumption of unpaid adultcare services provided
I8	Voluntary activity	Full Market-equivalent intermediate consumption of unpaid volunteering services provided
E1	Education	England attainment index (Secondary school)
E2	Education	Wales attainment index (Secondary school)
E3	Education	Scotland attainment index
E4	Education	NI attainment index
E5	Education	England attainment index (Primary school)
E6	Education	Wales attainment index (Primary school)
E7	Education	Scotland attainment index (Primary school)
E8	Education	NI attainment index (primaru school)
E9	Education	Bullying - Final Index
E10	Education	ITT QA Index
E-AGG	Education	Final QA Index
POS1	Public Order and Safety	Reoffending Index
POS2	Public Order and Safety	Safety index
POS3	Public Order and Safety	Escape index
POS-AGG	Public Order and Safety	Final QA Index
E	Education	Final QA Index
ASC	Adult Social Care	Final QA Index
CSC	Children's Social Care	Final QA Index
POS	Public Order and Safety	Final QA Index

Table 1: Target variables description.

	O1	**O2**	**O3**	**O4**	**O5**	**O6**	**O7**	**O8**
From	2006	2006	2006	2006	2006	2006	2006	2006
To	2016	2016	2016	2016	2016	2016	2016	2016
Sample Size	11	11	11	11	11	11	11	11
Sample Mean	3.040	7.088	3.864	1.157	3.991	5.646	3.786	0.725
Sample SD	1.216	4.827	4.610	14.441	1.094	2.865	4.174	3.472
Min	1.209	-0.072	-3.855	-24.390	2.221	1.389	-1.316	-7.211
Max	5.244	13.434	11.538	23.285	5.793	9.074	10.649	5.673
Skewness	0.223	-0.187	0.149	-0.340	0.215	-0.275	0.354	-0.720
Kurtosis	-1.216	-1.754	-1.164	-1.143	-1.272	-1.655	-1.573	0.017
JB_Stat	0.446	1.107	0.332	0.542	0.491	1.049	1.049	1.459
JB_Pval	0.800	0.575	0.847	0.763	0.782	0.592	0.592	0.482

	O1	**O2**	**O3**	**O4**	**O5**	**O6**	**O7**
From	2006	2006	2006	2006	2006	2006	2006
To	2016	2016	2016	2016	2016	2016	2016
Sample Size	11.000	11.000	11.000	11.000	11.000	11.000	11.000
Sample Mean	3.226	2.296	2.975	2.654	2.400	2.542	1.756
Sample SD	1.132	5.605	2.050	14.972	3.130	3.203	4.661
Min	1.362	-6.390	-1.369	-24.331	-1.889	-1.793	-6.872
Max	4.451	10.072	5.659	23.285	8.371	9.495	9.439
Skewness	-0.463	-0.048	-0.529	-0.436	0.123	0.516	-0.355
Kurtosis	-1.653	-1.426	-0.610	-1.149	-0.974	-0.382	-0.742
JB_Stat	1.383	0.556	0.689	0.728	0.175	0.662	0.341
JB_Pval	0.501	0.757	0.709	0.695	0.916	0.718	0.843

Table 2: Descriptive statistics for UHP target variables.

	E1	**E2**	**E3**	**E4**	**E5**	**E6**	**E7**	**E8**	**E9**
From	2006	2006	2006	2006	2006	2006	2006	2006	2006
To	2019	2019	2019	2019	2019	2019	2019	2019	2019
Sample Size	14	14	14	14	14	14	14	14	14
Sample Mean	1.895	1.318	1.167	2.780	1.702	0.939	-0.311	0.731	-0.195
Sample SD	2.241	2.852	1.691	1.788	2.100	1.240	2.615	1.618	1.831
Min	-1.075	-5.601	-3.702	0.000	-1.761	-1.350	-4.576	-3.758	-2.836
Max	6.505	6.367	2.954	6.821	6.321	2.403	5.839	2.853	2.772
Skewness	0.493	-0.572	-1.581	0.735	0.450	-0.683	0.575	-1.474	0.100
Kurtosis	-1.109	0.363	2.169	-0.222	-0.447	-0.985	0.072	1.670	-1.170
JB_Stat	1.090	1.428	12.521	1.604	0.590	1.618	1.148	9.739	0.478
JB_Pval	0.580	0.490	0.002	0.448	0.745	0.445	0.563	0.008	0.787
	E10	**POS1**	**POS2**	**POS3**	**E**	**ASC**	**CSC**	**POS**	
From	2006	2006	2006	2006	2006	2006	2006	2006	
To	2019	2019	2019	2019	2019	2019	2019	2019	
Sample Size	14	14	14	14	14	14	14	14	
Sample Mean	0.396	-0.986	-3.748	0.997	1.228	0.163	1.917	-2.435	
Sample SD	1.425	4.488	14.053	2.299	0.667	1.102	3.635	3.385	
Min	-2.645	-7.714	-23.701	-2.920	-0.086	-1.565	-2.369	-6.217	
Max	3.986	7.124	20.184	4.608	2.530	1.810	11.969	4.821	
Skewness	0.459	0.231	0.127	-0.031	0.381	0.155	1.339	0.912	
Kurtosis	1.402	-1.109	-1.491	-1.278	0.027	-1.518	1.525	-0.483	
JB_Stat	3.200	0.535	0.959	0.590	0.575	1.027	8.173	2.428	
JB_Pval	0.202	0.765	0.619	0.745	0.750	0.598	0.017	0.297	

Table 3: Descriptive statistics for PSP target variables.

	O1	O2	O3	O4	O5	O6	O7	O8	O-AGG	I1	I2	I3	I4	I5	I6	I7	I-AGG
AR(1)	6,439.252	14,546.573	14,136.546	1,536.421	945.710	11,687.119	3,950.996	1,226.739	25,411.121	1,942.953	3,179.226	4,055.950	297.626	826.409	1,928.369	551.505	7,395.048
AR(P)	0.938	0.990	1.001	0.709	1.039	1.019	0.880	0.921	0.988	1.020	0.490	1.038	0.953	0.979	0.987	1.039	1.013
AVG-AR	0.958	0.993	1.001	0.795	1.011	1.011	0.917	0.944	0.992	1.009	0.630	1.013	0.965	0.985	0.991	1.024	1.009
BFS-L	2.411	1.485	1.909	1.040	1.764	0.913	1.037	1.620	1.540	1.408	1.918	1.226	1.585	0.495	0.596	2.042	1.413
Ridge-L	0.937	0.986	0.983	0.784	1.103	0.923	0.899	0.994	0.995	1.150	0.810	1.041	1.028	0.980	0.988	0.993	0.968
Lasso-L	0.753	1.091	0.994	0.926	1.521	1.079	0.882	0.997	0.979	0.994	0.917	1.041	0.979	0.978	0.956	0.985	1.079
EN05-L	0.895	0.987	1.003	0.855	1.362	1.054	0.860	1.001	0.898	1.013	1.008	1.040	0.980	0.979	0.967	0.986	1.028
AdRidge-L	0.750	1.129	1.078	0.959	1.692	0.950	0.928	1.218	0.845	1.516	1.404	1.149	1.155	0.824	1.063	1.084	0.952
AdLasso-L	1.016	1.233	1.331	0.924	1.663	1.092	0.880	0.998	0.893	1.062	0.894	1.052	0.993	0.967	0.931	1.028	1.081
AdEN05-L	0.721	1.130	1.147	0.894	1.724	1.098	0.857	1.016	0.858	1.039	0.908	1.052	1.146	0.967	0.920	1.086	1.084
PCA(1)-L	0.821	1.223	0.694	0.884	1.090	0.975	0.915	1.065	1.050	1.090	0.756	1.048	1.137	0.954	0.979	0.793	0.873
PCA(2)-L	0.781	1.253	0.721	0.884	1.072	1.049	0.933	1.074	0.987	1.071	0.813	1.029	1.143	0.977	1.012	0.908	0.875
PCA(3)-L	0.798	1.600	0.671	0.871	1.307	0.915	0.913	1.117	1.104	1.329	0.762	1.004	1.094	0.946	0.958	0.895	0.851
PLS(1)-L	0.821	1.223	0.694	0.884	1.090	0.975	0.915	1.065	1.050	1.090	0.756	1.048	1.137	0.954	0.979	0.793	0.873
PLS(2)-L	0.781	1.253	0.721	0.884	1.072	1.049	0.933	1.074	0.987	1.071	0.813	1.029	1.143	0.977	1.012	0.908	0.875
PLS(3)-L	0.798	1.600	0.671	0.871	1.307	0.915	0.913	1.117	1.104	1.329	0.762	1.004	1.094	0.946	0.958	0.895	0.851
RF-L	0.900	0.963	0.939	0.828	1.096	0.939	0.884	1.019	0.899	1.222	0.885	1.013	1.032	1.003	1.025	1.040	0.972
MLP-L	1.243	2.246	0.962	0.670	1.161	1.207	1.400	1.117	1.241	1.279	1.391	1.320	1.119	1.645	1.144	1.269	2.014
EML-L	0.905	1.109	0.885	0.728	1.024	1.070	0.855	1.002	1.047	1.197	1.044	1.037	1.028	0.983	0.992	1.169	0.993
SVR-L	0.964	1.079	0.942	0.808	1.116	1.020	0.860	1.022	0.999	1.084	0.989	1.077	1.004	1.001	1.015	0.999	1.054
AVG-BFS-PEN-L	0.750	1.042	1.112	0.899	1.443	0.965	0.780	1.114	0.959	1.125	0.966	1.058	1.098	0.836	0.912	1.094	1.021
AVG-PCA-L	0.796	1.351	0.679	0.879	1.152	0.974	0.919	1.085	1.039	1.145	0.761	1.026	1.123	0.958	0.982	0.863	0.865
AVG-PLS-L	0.796	1.351	0.679	0.879	1.152	0.974	0.919	1.085	1.039	1.145	0.761	1.026	1.123	0.958	0.982	0.863	0.865
AVG-NL-L	0.986	1.281	0.842	0.751	1.032	1.041	0.959	1.037	1.022	1.171	1.034	1.009	1.028	1.115	1.034	1.053	1.245
AVG-ALL-L	0.811	1.237	0.794	0.848	1.183	0.978	0.889	1.078	0.978	1.128	0.852	1.017	1.086	0.962	0.973	0.945	0.991
AVG-ALL-wAR-L	0.877	1.109	0.886	0.816	1.094	0.989	0.902	1.008	0.980	1.063	0.738	1.013	1.019	0.973	0.981	0.979	0.998
BFS-S	0.839	1.452	1.223	1.342	2.471	0.646	1.130	1.454	0.710	1.161	1.703	1.306	1.957	0.600	0.701	2.755	1.919
Ridge-S	0.917	1.248	0.909	0.782	1.158	1.023	0.875	1.023	0.971	0.997	0.743	1.044	1.004	1.018	1.010	1.007	0.921
Lasso-S	0.986	0.871	1.036	0.787	1.013	0.996	1.001	0.951	1.091	0.902	0.904	1.041	1.045	0.806	0.803	0.954	1.023
EN05-S	0.942	0.870	1.016	0.785	1.025	0.992	0.916	0.963	1.078	0.919	0.936	1.040	0.995	0.823	0.867	0.921	0.990
AdRidge-S	0.940	1.197	1.256	0.944	1.134	1.185	1.027	0.929	1.105	1.435	0.552	1.112	1.255	0.690	0.735	0.951	0.849
AdLasso-S	0.957	0.929	1.192	0.829	0.951	1.304	1.152	0.979	1.137	0.878	0.874	1.052	1.148	0.763	0.654	0.949	1.065
AdEN05-S	0.932	0.907	1.160	0.822	0.950	1.304	1.090	0.983	1.165	0.939	0.852	1.052	1.123	0.765	0.719	0.951	1.074
PCA(1)-S	0.815	1.140	0.738	0.836	1.087	1.168	0.853	1.035	0.819	0.999	0.911	0.969	1.086	1.013	1.030	1.051	0.735
PCA(2)-S	1.024	1.442	0.697	0.889	1.077	1.202	0.991	1.096	1.076	1.553	0.638	1.112	1.071	1.017	1.058	0.785	1.141
PCA(3)-S	1.020	1.551	0.980	0.954	1.311	0.731	1.112	1.119	1.154	1.765	0.657	1.146	1.124	1.138	1.118	0.878	1.148
PLS(1)-S	0.815	1.140	0.738	0.836	1.087	1.168	0.853	1.035	0.819	0.999	0.911	0.969	1.086	1.013	1.030	1.051	0.735
PLS(2)-S	1.024	1.442	0.697	0.889	1.077	1.202	0.991	1.096	1.076	1.553	0.638	1.112	1.071	1.017	1.058	0.785	1.141
PLS(3)-S	1.020	1.551	0.980	0.954	1.311	0.731	1.112	1.119	1.154	1.765	0.657	1.146	1.124	1.138	1.118	0.878	1.148
RF-S	0.907	1.101	0.908	0.820	1.099	0.965	0.891	1.004	0.980	1.230	0.769	1.024	1.014	0.960	0.982	1.025	0.978
MLP-S	1.210	2.227	1.380	0.674	1.169	1.204	1.411	1.112	1.216	1.112	1.402	1.329	1.114	1.758	1.138	1.405	1.986
EML-S	0.905	1.102	2.272	0.728	1.014	1.070	0.855	1.002	1.047	1.099	1.044	1.037	1.028	0.983	0.992	1.141	0.993
SVR-S	0.977	1.133	0.962	0.822	1.151	1.040	0.877	1.026	1.014	1.067	0.906	1.069	1.026	0.980	0.989	0.979	1.009
AVG-BFS-PEN-S	0.802	0.928	0.994	0.868	1.092	1.029	0.984	1.023	1.002	0.895	0.383	1.069	1.172	0.757	0.759	1.091	1.006
AVG-PCA-S	0.907	1.330	0.760	0.891	1.137	1.019	0.974	1.081	1.006	1.291	0.475	1.068	1.088	1.055	1.067	0.881	0.968
AVG-PLS-S	0.907	1.330	0.760	0.891	1.137	1.019	0.974	1.081	1.006	1.291	0.475	1.068	1.088	1.055	1.067	0.881	0.968
AVG-NL-S	0.986	1.328	1.190	0.754	1.052	1.056	0.967	1.033	1.037	1.110	0.984	1.009	1.028	1.134	1.013	1.069	1.230
AVG-ALL-S	0.898	1.185	0.806	0.847	1.090	1.019	0.958	1.044	0.979	1.142	0.530	1.030	1.084	0.995	0.972	0.941	1.011
AVG-ALL-wAR-S	0.924	1.081	0.895	0.817	1.047	1.009	0.936	0.993	0.979	1.072	0.574	1.019	1.019	0.990	0.981	0.970	0.994

Table 4: UHP nowcasting cross-validation results from 2013 to 2016 using MV1 pattern for missing values. Actual values are reported for AR(1). Relative values to the AR(1) benchmark are reported for all other models.

	O1	O2	O3	O4	O5	O6	O7	O8	O-AGG	I1	I2	I3	I4	I5	I6	I7	I-AGG
AR(1)	6,439.252	14,546.573	14,136.546	1,536.421	945.710	11,687.119	3,950.996	1,226.739	25,411.121	1,942.953	3,179.226	4,055.950	297.626	826.409	1,928.369	551.505	7,395.048
AR(P)	0.938	0.990	1.001	0.709	1.039	1.019	0.880	0.921	0.988	1.020	0.490	1.038	0.953	0.979	0.987	1.039	1.013
AVG-AR	0.958	0.993	1.001	0.795	1.011	1.011	0.917	0.944	0.992	1.009	0.630	1.013	0.965	0.985	0.991	1.024	1.009
BFS-L	1.006	0.770	1.794	0.931	1.995	1.054	1.009	1.615	1.443	1.692	1.865	0.877	2.140	0.474	0.597	2.055	1.435
Ridge-L	0.939	0.984	0.984	0.785	1.110	1.013	0.911	0.996	0.999	1.154	0.808	1.043	0.969	0.980	0.987	0.996	0.999
Lasso-L	0.985	0.896	1.001	0.814	1.536	1.074	0.898	0.995	1.024	0.996	0.985	1.087	0.953	0.978	0.972	0.994	1.017
EN05-L	1.014	0.961	0.983	0.790	1.287	1.080	0.857	0.995	0.963	1.088	1.008	1.040	0.994	0.978	0.980	0.909	1.032
AdRidge-L	1.028	1.010	1.119	0.908	1.557	1.096	1.040	1.178	1.042	1.341	1.572	1.133	1.250	0.841	1.042	0.903	1.033
AdLasso-L	0.914	0.874	1.155	0.845	1.766	1.111	0.920	0.997	1.055	1.039	0.989	1.120	0.953	0.965	0.949	0.945	1.103
AdEN05-L	1.057	0.978	1.028	0.810	1.735	1.124	0.916	0.993	1.035	1.088	0.994	1.067	1.003	0.965	0.959	0.935	1.103
PCA(1)-L	0.816	1.245	0.709	0.907	1.027	1.037	0.968	1.086	0.990	1.166	0.737	1.074	1.166	0.973	1.011	0.815	0.987
PCA(2)-L	0.809	1.284	0.710	0.920	1.045	0.999	0.985	1.091	1.016	1.135	0.844	1.086	1.181	0.969	1.010	0.828	1.024
PCA(3)-L	1.078	1.468	0.673	0.929	1.115	0.832	1.121	1.161	1.164	1.379	0.634	1.142	1.078	1.024	1.015	0.798	1.063
PLS(1)-L	0.816	1.245	0.709	0.907	1.027	1.037	0.968	1.086	0.990	1.166	0.737	1.074	1.166	0.973	1.011	0.815	0.987
PLS(2)-L	0.809	1.284	0.710	0.920	1.045	0.999	0.985	1.091	1.016	1.135	0.844	1.086	1.181	0.969	1.010	0.828	1.024
PLS(3)-L	1.078	1.468	0.673	0.929	1.115	0.832	1.121	1.161	1.164	1.379	0.634	1.142	1.078	1.024	1.015	0.798	1.063
RF-L	0.951	1.038	0.950	0.819	1.063	1.004	0.903	1.004	0.925	1.181	0.891	1.023	1.033	1.005	1.006	1.032	0.988
MLP-L	1.240	2.288	1.058	0.664	1.081	1.202	1.353	1.098	1.753	1.265	1.367	1.288	1.132	1.771	1.141	1.431	1.960
EML-L	0.905	1.084	0.863	0.729	1.014	1.070	0.855	1.002	1.047	1.195	1.044	1.037	1.030	0.983	0.992	1.141	0.993
SVR-L	0.969	1.083	0.951	0.802	1.113	1.034	0.873	1.009	1.011	1.074	1.006	1.081	0.992	0.993	1.002	1.005	1.067
AVG-BFS-PEN-L	0.973	0.886	1.050	0.812	1.474	1.068	0.817	1.105	0.949	1.144	1.017	1.021	1.043	0.839	0.908	1.002	1.053
AVG-PCA-L	0.892	1.303	0.695	0.917	1.040	0.945	1.016	1.112	1.051	1.180	0.715	1.099	1.140	0.987	1.010	0.811	1.020
AVG-PLS-L	0.892	1.303	0.695	0.917	1.040	0.945	1.016	1.112	1.051	1.180	0.715	1.099	1.140	0.987	1.010	0.811	1.020
AVG-NL-L	1.000	1.300	0.857	0.745	1.006	1.060	0.953	1.025	1.155	1.149	1.027	1.008	1.030	1.151	1.025	1.076	1.241
AVG-ALL-L	0.938	1.183	0.803	0.843	1.127	0.998	0.944	1.087	1.011	1.152	0.847	1.046	1.079	0.988	0.984	0.899	1.072
AVG-ALL-wAR-L	0.944	1.080	0.891	0.814	1.067	1.001	0.930	1.013	0.998	1.073	0.736	1.026	1.015	0.986	0.987	0.956	1.038
BFS-S	1.489	1.179	1.227	1.197	2.495	0.696	1.585	1.522	0.826	2.601	1.730	1.188	1.929	0.921	0.718	2.758	1.635
Ridge-S	0.952	1.251	0.903	0.781	1.242	1.021	0.894	1.008	1.000	1.020	0.773	1.045	0.987	0.986	0.977	1.001	0.948
Lasso-S	1.256	0.864	1.040	0.799	1.039	0.983	1.093	0.953	1.120	1.148	0.915	1.086	1.063	0.825	0.816	0.985	1.016
EN05-S	1.009	0.900	1.010	0.792	1.039	1.000	0.963	0.955	1.176	1.060	0.866	1.071	0.976	0.845	0.870	0.981	1.005
AdRidge-S	1.229	0.907	1.022	0.943	1.343	1.195	1.086	0.919	1.164	1.405	0.579	1.179	1.272	0.713	0.750	0.998	0.822
AdLasso-S	1.309	0.944	1.190	0.899	1.039	1.303	1.221	0.985	1.153	1.613	0.894	1.120	1.124	0.800	0.671	1.001	1.003
AdEN05-S	1.219	0.852	1.153	0.899	1.039	0.942	1.170	0.970	1.148	1.419	0.857	1.118	1.136	0.792	0.707	1.008	1.145
PCA(1)-S	0.834	1.188	0.690	0.822	1.099	1.055	0.867	1.005	0.929	1.024	0.970	0.982	1.064	0.985	1.003	0.951	0.750
PCA(2)-S	0.964	1.480	0.752	0.892	1.047	1.143	1.042	1.084	1.225	1.720	0.724	1.156	1.083	0.995	1.042	0.685	1.293
PCA(3)-S	1.283	1.444	1.071	0.944	1.552	0.790	1.078	1.113	1.194	1.692	0.956	1.120	1.146	1.118	1.040	0.707	1.087
PLS(1)-S	0.834	1.188	0.690	0.822	1.099	1.055	0.867	1.005	0.929	1.024	0.970	0.982	1.064	0.985	1.003	0.951	0.750
PLS(2)-S	0.964	1.480	0.752	0.892	1.047	1.143	1.042	1.084	1.225	1.720	0.724	1.156	1.083	0.995	1.042	0.685	1.293
PLS(3)-S	1.283	1.444	1.071	0.944	1.552	0.790	1.078	1.113	1.194	1.692	0.956	1.120	1.146	1.118	1.040	0.707	1.087
RF-S	0.873	1.091	0.878	0.825	1.064	0.947	0.907	0.996	1.038	1.282	0.788	1.001	1.018	0.915	0.948	1.013	0.968
MLP-S	1.236	2.164	1.381	0.669	1.295	1.224	1.368	1.109	1.235	1.066	1.246	1.310	1.129	1.772	1.144	1.306	1.976
EML-S	0.905	1.110	2.252	0.729	1.014	1.070	0.855	1.002	1.047	1.099	1.090	1.037	1.027	0.983	0.992	1.170	0.993
SVR-S	0.995	1.126	0.970	0.811	1.154	1.050	0.888	1.008	1.025	1.047	0.936	1.069	1.007	0.972	0.976	0.986	1.033
AVG-BFS-PEN-S	1.129	0.933	0.993	0.888	1.167	0.981	1.132	1.031	1.030	1.149	0.402	1.098	1.168	0.804	0.755	1.176	1.010
AVG-PCA-S	0.867	1.305	0.743	0.881	1.178	0.917	0.916	1.064	1.107	1.172	0.513	1.075	1.091	1.016	1.018	0.749	0.995
AVG-PLS-S	0.867	1.305	0.743	0.881	1.178	0.917	0.916	1.064	1.107	1.172	0.513	1.075	1.091	1.016	1.018	0.749	0.995
AVG-NL-S	0.983	1.301	1.198	0.750	1.069	1.059	0.961	1.026	1.064	1.107	0.960	1.011	1.029	1.124	1.004	1.044	1.232
AVG-ALL-S	0.939	1.169	0.793	0.846	1.133	0.962	0.967	1.038	1.047	1.120	0.494	1.050	1.084	0.980	0.944	0.905	1.031
AVG-ALL-wAR-S	0.947	1.076	0.878	0.817	1.070	0.983	0.941	0.989	1.014	1.060	0.558	1.029	1.019	0.983	0.967	0.950	1.010

Table 5: UHP nowcasting cross-validation results from 2013 to 2016 using MV2 pattern for missing values. Actual values are reported for AR(1). Relative values to the AR(1) benchmark are reported for all other models.

	O1	O2	O3	O4	O5	O6	O7	O8	O-AGG	I1	I2	I3	I4	I5	I6	I7	I-AGG
AR(1)	6,439.252	14,546.573	14,136.546	1,536.421	945.710	11,687.119	3,950.996	1,226.739	25,411.121	1,942.953	3,179.226	4,055.950	297.626	826.409	1,928.369	551.505	7,395.048
AR(P)	0.938	0.990	1.001	0.709	1.039	1.019	0.880	0.921	0.988	1.020	0.490	1.038	0.953	0.979	0.987	1.039	1.013
AVG-AR	0.958	0.993	1.001	0.795	1.011	1.011	0.917	0.944	0.992	1.009	0.630	1.013	0.965	0.985	0.991	1.024	1.009
BFS-L	1.137	4.771	2.382	1.239	1.923	0.944	1.086	1.608	0.630	1.099	3.185	1.406	1.779	0.777	1.453	2.604	1.299
Ridge-L	1.031	0.986	0.976	0.834	1.151	1.018	0.977	0.995	0.999	1.051	0.837	1.092	1.017	0.995	1.010	1.000	0.982
Lasso-L	1.071	0.755	0.988	0.781	1.285	1.162	1.066	1.017	1.427	0.997	1.181	1.038	0.953	0.811	0.832	1.006	1.024
EN05-L	1.087	0.846	0.993	0.851	1.181	1.172	0.999	1.002	1.225	1.020	1.007	1.038	0.970	0.905	0.897	0.977	1.017
AdRidge-L	0.936	0.755	0.837	0.876	1.318	1.110	1.077	1.239	1.417	0.941	1.124	1.214	1.145	0.746	0.859	0.978	1.395
AdLasso-L	0.989	0.629	0.984	0.780	1.263	1.135	1.127	1.048	1.564	0.821	1.279	1.038	0.953	0.759	0.776	1.060	1.224
AdEN05-L	1.145	0.628	0.983	0.883	1.496	1.140	0.985	1.046	1.534	1.025	1.183	1.038	0.980	0.806	0.809	0.949	1.150
PCA(1)-L	0.912	1.151	0.756	1.040	1.489	1.178	1.025	1.126	0.889	1.496	0.886	1.137	1.275	0.946	0.970	1.035	1.166
PCA(2)-L	1.345	1.807	0.986	1.055	1.252	1.239	1.347	1.146	1.234	0.945	0.530	1.309	1.299	1.180	1.268	0.924	1.139
PCA(3)-L	1.771	2.317	0.964	1.072	1.511	0.982	1.679	1.235	1.586	1.804	0.791	1.453	1.100	1.323	1.328	1.006	1.284
PLS(1)-L	0.912	1.151	0.756	1.040	1.489	1.178	1.025	1.126	0.889	1.496	0.886	1.137	1.275	0.946	0.970	1.035	1.166
PLS(2)-L	1.345	1.807	0.986	1.055	1.252	1.239	1.347	1.146	1.234	0.945	0.530	1.309	1.299	1.180	1.268	0.924	1.139
PLS(3)-L	1.771	2.317	0.964	1.072	1.511	0.982	1.679	1.235	1.586	1.804	0.791	1.453	1.100	1.323	1.328	1.006	1.284
RF-L	0.998	1.092	0.944	0.858	1.071	0.985	0.944	1.013	1.012	1.078	0.847	1.078	1.016	0.976	0.994	1.008	0.940
MLP-L	1.214	2.255	1.298	0.664	1.183	1.215	1.359	1.082	1.210	1.120	1.538	1.314	1.120	1.748	1.139	1.139	1.960
EML-L	0.905	1.070	0.885	0.729	1.024	1.070	0.855	1.002	1.047	1.099	1.044	1.037	1.029	0.983	0.992	1.104	0.993
SVR-L	0.996	1.083	0.961	0.796	1.095	1.049	0.895	0.986	1.045	1.031	1.033	1.090	0.986	0.985	0.990	1.003	1.067
AVG-BFS-PEN-L	1.036	1.065	0.942	0.873	1.232	1.056	1.041	1.113	1.082	0.892	1.200	1.082	1.064	0.803	0.916	0.855	1.136
AVG-PCA-L	1.307	1.677	0.882	1.051	1.336	1.101	1.333	1.155	1.206	1.336	0.444	1.275	1.219	1.131	1.167	0.954	1.134
AVG-PLS-L	1.307	1.677	0.882	1.051	1.336	1.101	1.333	1.155	1.206	1.336	0.444	1.275	1.219	1.131	1.167	0.954	1.134
AVG-NL-L	1.014	1.292	0.851	0.753	1.040	1.062	0.973	1.014	1.060	1.068	1.070	1.046	1.020	1.139	1.019	0.971	1.229
AVG-ALL-L	1.143	1.384	0.862	0.925	1.206	1.064	1.157	1.091	1.087	1.085	0.743	1.136	1.119	1.043	1.050	0.829	1.117
AVG-ALL-wAR-L	1.034	1.179	0.906	0.853	1.103	1.030	1.032	1.010	1.033	1.025	0.685	1.071	1.035	1.008	1.015	0.919	1.051
BFS-S	1.697	3.453	1.407	1.044	2.762	1.158	1.077	1.607	0.664	2.373	1.822	1.362	1.087	0.788	1.157	2.222	1.687
Ridge-S	0.996	1.322	0.909	0.782	1.302	1.018	0.918	1.000	1.072	1.021	0.749	1.055	0.973	1.014	0.983	0.982	0.908
Lasso-S	0.928	1.307	0.990	0.785	1.039	0.969	1.005	1.019	1.396	1.063	0.874	1.040	1.091	0.782	0.821	0.956	1.072
EN05-S	0.940	1.480	1.005	0.783	1.089	1.014	0.972	0.976	1.425	1.085	0.867	1.033	0.955	0.871	0.922	0.957	1.056
AdRidge-S	0.930	1.675	0.541	0.841	1.440	1.195	1.026	0.927	1.416	1.679	0.854	0.944	1.072	0.759	0.755	0.951	0.828
AdLasso-S	0.929	1.527	0.992	0.806	1.039	1.304	1.036	1.059	1.400	1.354	0.820	1.037	1.183	0.693	0.719	0.948	0.712
AdEN05-S	0.912	1.602	0.998	0.806	1.329	0.970	1.056	1.006	1.409	1.101	1.181	1.046	1.072	0.706	0.734	0.958	1.245
PCA(1)-S	0.861	1.180	0.731	0.892	1.131	1.117	0.940	1.084	0.803	1.015	1.094	1.065	1.121	1.021	1.042	1.039	1.081
PCA(2)-S	1.125	1.609	0.570	0.883	1.211	1.118	1.022	1.098	1.189	1.640	0.741	1.133	1.024	1.105	1.099	0.792	1.090
PCA(3)-S	1.415	1.905	0.949	0.984	2.028	0.563	1.125	1.063	1.286	2.388	0.787	1.084	1.140	1.132	1.042	0.728	0.804
PLS(1)-S	0.861	1.180	0.731	0.892	1.131	1.117	0.940	1.084	0.803	1.015	1.094	1.065	1.121	1.021	1.042	1.039	1.081
PLS(2)-S	1.125	1.609	0.570	0.883	1.211	1.118	1.022	1.098	1.189	1.640	0.741	1.133	1.024	1.105	1.099	0.792	1.090
PLS(3)-S	1.415	1.905	0.949	0.984	2.028	0.563	1.125	1.063	1.286	2.388	0.787	1.084	1.140	1.132	1.042	0.728	0.804
RF-S	0.970	1.151	0.842	0.803	1.054	0.990	0.882	1.009	1.060	1.190	0.838	1.001	0.977	0.947	0.949	0.989	0.940
MLP-S	1.217	2.294	1.336	0.645	1.262	1.213	1.409	1.114	1.241	1.157	1.544	1.331	1.144	1.751	1.137	1.146	1.959
EML-S	0.905	1.074	2.244	0.729	1.014	1.070	0.855	1.002	1.047	1.099	1.044	1.037	1.029	0.983	0.992	1.060	0.993
SVR-S	0.991	1.109	0.967	0.801	1.163	1.044	0.869	0.983	1.034	1.043	0.968	1.060	0.995	0.981	0.980	0.991	1.021
AVG-BFS-PEN-S	0.979	1.646	0.848	0.828	1.336	1.039	0.989	1.068	1.234	1.248	0.981	1.055	0.974	0.780	0.843	0.925	0.989
AVG-PCA-S	1.038	1.471	0.674	0.913	1.428	0.870	0.990	1.072	1.088	1.349	0.457	1.080	1.079	1.079	1.057	0.803	0.914
AVG-PLS-S	1.038	1.471	0.674	0.913	1.428	0.870	0.990	1.072	1.088	1.349	0.457	1.080	1.079	1.079	1.057	0.803	0.914
AVG-NL-S	1.004	1.314	1.169	0.736	1.065	1.066	0.960	1.023	1.075	1.108	1.047	1.020	1.021	1.129	1.003	0.945	1.215
AVG-ALL-S	1.005	1.418	0.698	0.843	1.299	0.954	0.960	1.049	1.094	1.235	0.631	1.035	1.028	1.007	0.983	0.775	0.978
AVG-ALL-wAR-S	0.981	1.198	0.837	0.815	1.151	0.980	0.936	0.995	1.038	1.105	0.620	1.023	0.994	0.994	0.986	0.898	0.985

Table 6: UHP nowcasting cross-validation results from 2013 to 2016 using MV3 pattern for missing values. Actual values are reported for AR(1). Relative values to the AR(1) benchmark are reported for all other models.

Variable	2013	2014	2015	2016	Variable	2013	2014	2015	2016
bp_gbcoalconsej	-0.191	-0.015	0.019	-0.086	**gbbank02072**	NA	NA	0.037	-0.065
bp_gbgasconscm	-0.077	-0.05	-0.029	0.058	**gbnaac22382**	NA	NA	NA	-0.196
bp_gboilcons	-0.194	NA	-0.165	0.177	**gbnaac00352**	NA	NA	0.137	-0.159
bp_gbenergyconsej	-0.192	-0.068	-0.072	0.079	**gbnaac00332**	NA	NA	NA	0.078
gbdemo2064	NA	NA	0.158	-0.205	**gbnaac00402**	-0.122	-0.023	-0.002	NA
gbdemo1287	-0.119	-0.018	NA	NA	**gbnaac00422**	NA	NA	NA	-0.034
gbdemo1180	NA	NA	-0.071	0.118	**gbnaac1098**	NA	NA	-0.118	NA
gbdemo0236	0.068	NA	0.127	-0.113	**gbcaes0167**	-0.258	-0.152	-0.134	0.096
gbdemo0234	-0.174	-0.175	-0.093	0.052	**wocaes0482**	-0.003	0.056	0.092	-0.16
gbdemo0233	NA	-0.122	NA	0.195	**ukepunnewsindex**	-0.021	0.04	0.064	NA
gbdemo0230	NA	0.089	0.083	-0.101	**gbepunnewsindex**	-0.014	0.116	0.105	-0.006
gbdemo0235	NA	-0.048	-0.026	-0.062	**gbepugpr0030**	-0.093	-0.141	-0.169	0.15
gbdemo0232	NA	0.219	0.194	-0.15	**gbfofi0150**	-0.22	-0.12	-0.13	0.09
gbdemo0231	NA	0.05	0.032	-0.027	**gbfofi0103**	-0.011	0.099	0.128	-0.188
ndgain_ndgain_tot_gb	NA	-0.227	-0.18	0.084	**gblama1207**	NA	NA	0.166	-0.17
undp_gii_0152	-0.019	0.047	0.117	-0.181	**gblama1356**	0.194	0.173	NA	NA
undemo0014	0.008	-0.058	-0.043	0.059	**gbmost0100**	NA	NA	0.097	-0.129
wpfi_gb_s	0.11	0.184	0.161	-0.141	**gbmost0043**	-0.03	-0.078	-0.015	-0.002
cs_gw_mwpa0169	-0.093	-0.133	-0.118	0.103	**gbmost0049**	0.045	0.041	NA	NA
gb_swiid_abs_red	NA	0.194	NA	NA	**gbbank6836**	-0.12	-0.027	0.025	-0.058
gb_swiid_gini_disp	0.087	-0.06	-0.057	0.092	**gbpric00011**	0.006	0.134	0.113	-0.179
gb_swiid_rel_red	0.092	NA	0.139	-0.151	**gbpric00141**	NA	0.038	0.009	-0.055
uspoli1141	-0.201	-0.112	-0.061	-0.036	**gbpric00131**	0.229	NA	NA	NA
uspoli0211	0.016	-0.084	-0.065	0.102	**gbpric00181**	-0.006	0.1	0.115	-0.181
uspoli0583	0.086	0.04	0.041	-0.003	**gbpric00161**	NA	0.096	0.135	-0.186
uspoli0955	-0.029	-0.142	-0.117	0.125	**gbpric00211**	NA	NA	-0.065	0.087
uspoli1513	0.124	0.05	0.014	0.027	**gbprod0515**	-0.199	-0.214	-0.172	0.17
uspoli0769	-0.047	-0.058	-0.062	0.121	**gbsurv0559**	NA	NA	-0.215	0.153
uspoli1699	-0.156	-0.146	-0.069	0.005	**gbsurv9466**	NA	NA	-0.029	0.05
gb_rl_est_esg	-0.086	-0.135	-0.083	0.027	**csiigbp**	NA	NA	0.193	-0.132
gb_pv_est_esg	NA	NA	-0.141	0.059	**gbcaes0058**	-0.114	0.035	NA	-0.153
gbinea0207	NA	NA	0.188	-0.174	**defra01**	0.216	0.231	0.191	-0.138
gbinea0048	NA	-0.144	-0.089	0.051	**gbeqin0057_pe**	NA	NA	-0.117	0.136
gbbopa00482	0.112	0.149	0.156	-0.135	**lse_all_mcap**	NA	-0.236	-0.201	0.14
gbbank02092	0.231	0.213	0.208	-0.201	**agbp10y**	0.181	0.118	0.146	-0.15
gbpric0009	0.055	NA	NA	NA	**ftseftp_bg05_avepr**	NA	0.187	0.091	-0.009
gbcons7097	-0.231	-0.191	-0.137	0.074	**gbpeur**	0.248	0.193	0.188	-0.153
gbwui	NA	NA	0.093	0.012	**gbp**	0.231	0.19	0.128	-0.03
gbbank02082	-0.195	-0.162	-0.13	0.097	**gb5ygov**	-0.211	-0.241	-0.178	0.109

Table 7: PCA(1)-S loadings using MV3 missing values pattern for O1 target.

	2013	2014	2015	2016
gbdemo0236	NA	NA	-0.804	NA
ndgain_vul_1143	-1.996	NA	NA	NA
wpfi_gb_r	NA	NA	NA	-0.163
uspoli0955	NA	NA	NA	0.007
gb_rl_est_esg	-0.107	-0.149	-0.144	NA
gbbopa00482	-0.01	NA	NA	-0.014
gbbank02082	NA	NA	NA	-0.299
gbmost0352	0	NA	NA	NA
gbmost0984	NA	NA	0.117	NA
gbpric00181	NA	-0.749	NA	1.58
gbpric00231	NA	-0.316	NA	NA
agbp10y	NA	0.689	NA	NA
gbpeur	NA	NA	0.077	NA

Table 8: BFS-L beta coefficient estimates using MV3 missing values pattern for O-AGG target.

Target	Top Model (MV3)	2017	2018	2019	2020	2021	2022
O1	PCA(1)-S	501,162	516,066	532,175	550,852	560,834	590,417
O2	AdEN05-L	448,451	506,379	564,392	613,041	663,344	712,072
O3	AdRidge-S	325,536	340,739	359,606	380,814	399,368	448,573
O4	MLP-S	4,963	4,610	5,084	4,628	5,178	4,709
O5	AR(1)	111,728	116,240	120,940	125,831	130,920	136,215
O6	PCA(3)-S	420,585	443,624	466,479	469,104	453,468	460,834
O7	EML-L	70,554	72,898	75,319	77,822	80,407	83,078
O8	AR(P)	23,934	23,934	23,934	23,934	23,934	23,934
AGG	BFS-L	1,890,609	2,015,403	2,142,847	2,239,181	2,344,118	2,583,486
I1	AdLasso-L	296,717	303,710	313,059	322,942	333,035	343,469
I2	AVG-PCA-L	56,002	57,381	57,902	57,200	59,040	62,756
I3	AdRidge-S	145,117	150,945	157,624	162,384	165,619	172,520
I4	Lasso-L	1,059	1,087	1,116	1,146	1,177	1,354
I5	AdLasso-S	19,060	19,476	19,497	19,888	20,099	20,308
I6	AdLasso-S	45,010	46,313	47,030	48,070	48,772	49,511
I7	PCA(3)-S	8,909	9,055	9,256	9,734	9,884	10,806
AGG	AdLasso-S	572,532	589,346	607,574	625,812	644,598	663,948

Table 9: UHP out-of-sample estimates for the top model of MV3 cross-validation. Illustrations for the above models are provided in the Appendix.

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E-AGG	POS1	POS2	POS3	POS-AGG
AR(1)	2.094	9.133	4.242	2.800	9.733	2.997	8.673	3.877	1.698	0.903	1.584	7.274	14.106	6.843	4.817
AR(P)	1.527	1.023	1.000	1.001	1.026	1.006	0.970	0.905	0.952	0.995	0.905	1.022	1.068	0.985	1.266
AVG-AR	1.308	1.014	1.000	1.001	1.017	1.004	0.980	0.933	0.966	0.957	0.936	1.013	1.044	0.990	1.176
BFS-L	0.960	1.000	1.140	1.906	0.827	29.758	1.008	1.357	1.034	2.112	1.349	0.869	1.211	0.970	1.469
Ridge-L	1.073	0.970	1.013	1.000	1.001	1.012	0.982	0.855	1.010	0.998	0.875	1.023	1.008	0.894	1.004
Lasso-L	0.421	1.019	1.005	1.115	1.022	1.161	0.928	1.076	1.075	1.346	0.791	0.996	1.054	0.840	0.985
EN05-L	0.641	1.020	1.004	1.077	1.004	1.045	0.923	1.041	1.068	1.013	0.803	1.017	1.067	0.772	0.905
AdRidge-L	0.473	0.978	0.977	1.451	1.083	1.040	0.926	1.106	1.103	1.488	0.840	1.078	1.016	0.743	0.857
AdLasso-L	0.390	1.002	1.045	1.353	1.022	1.236	0.923	1.094	1.123	1.684	0.801	0.988	1.079	0.833	0.979
AdEN05-L	0.398	1.002	1.043	1.360	1.010	1.192	0.914	1.052	1.133	1.422	0.764	1.059	0.997	0.727	0.861
PCA(1)-L	1.220	1.002	1.004	1.040	1.004	1.001	0.984	0.851	1.057	1.095	0.895	0.995	0.996	0.920	0.940
PCA(2)-L	0.938	0.995	1.013	1.057	1.004	1.062	0.981	0.883	1.006	1.004	0.896	0.963	0.837	0.868	0.843
PCA(3)-L	0.585	0.957	1.015	0.954	0.974	1.065	0.994	0.921	1.237	0.986	1.062	1.083	0.886	0.887	1.119
PLS(1)-L	1.220	1.002	1.004	1.040	1.004	1.001	0.984	0.851	1.057	1.095	0.895	0.995	0.996	0.920	0.940
PLS(2)-L	0.938	0.995	1.013	1.057	1.004	1.062	0.981	0.883	1.006	1.004	0.896	0.963	0.837	0.868	0.843
PLS(3)-L	0.585	0.957	1.015	0.954	0.974	1.065	0.994	0.921	1.237	0.986	1.062	1.083	0.886	0.887	1.119
RF-L	1.007	0.967	0.994	1.141	0.980	1.066	0.966	0.922	1.016	1.021	0.926	1.016	1.015	0.921	0.942
MLP-L	1.105	1.076	1.023	0.980	1.185	1.031	0.958	0.884	0.997	1.808	0.653	1.045	1.304	0.745	0.841
EML-L	0.839	1.027	1.002	0.988	0.984	1.027	1.005	0.848	1.022	1.002	0.778	0.986	0.927	0.895	0.901
SVR-L	1.071	0.995	1.000	1.049	1.003	1.024	0.979	0.872	1.069	0.949	0.889	1.036	0.974	0.987	0.931
AVG-BFS-PEN-L	0.574	0.995	1.022	1.309	0.987	29.758	0.943	1.072	1.040	1.282	0.880	0.995	1.044	0.743	0.997
AVG-PCA-L	0.909	0.984	1.010	1.001	0.994	1.042	0.986	0.884	1.094	1.022	0.950	1.010	0.905	0.891	0.965
AVG-PLS-L	0.909	0.984	1.010	1.001	0.994	1.042	0.986	0.884	1.094	1.022	0.950	1.010	0.905	0.891	0.965
AVG-NL-L	0.803	1.004	1.005	1.038	1.038	1.034	0.977	0.879	1.021	0.671	0.804	1.020	1.050	0.882	0.901
AVG-ALL-L	0.794	0.990	1.011	1.084	1.002	29.757	0.973	0.926	1.049	0.966	0.893	1.008	0.967	0.846	0.953
AVG-ALL-wAR-L	1.020	1.002	1.005	1.042	1.009	29.755	0.976	0.924	1.005	0.882	0.914	1.009	1.003	0.917	1.062
BFS-S	1.799	1.080	1.268	1.535	1.210	1.429	0.845	3.195	0.981	1.604	1.188	0.709	0.999	0.670	1.631
Ridge-S	1.168	1.028	0.997	1.001	1.001	1.006	0.982	0.878	1.006	0.998	0.853	1.007	0.961	0.861	0.976
Lasso-S	0.699	0.998	0.996	1.394	1.041	1.145	0.958	0.946	1.156	1.270	0.801	1.009	1.104	0.804	0.842
EN05-S	0.827	1.028	0.989	1.368	1.001	1.039	0.968	0.869	1.160	0.901	0.888	1.009	1.084	0.814	0.850
AdRidge-S	0.835	1.092	0.962	1.472	1.019	1.193	0.949	1.067	1.128	1.440	0.804	1.011	1.252	0.827	0.813
AdLasso-S	0.920	0.986	1.015	1.589	1.053	1.297	0.929	1.004	1.097	1.501	0.705	1.009	1.097	0.724	0.814
AdEN05-S	0.912	1.019	1.000	1.643	1.041	1.285	0.965	1.004	1.024	1.484	0.822	1.009	1.113	0.828	0.847
PCA(1)-S	0.976	1.035	0.996	1.062	1.026	1.003	0.975	0.911	0.967	0.964	0.851	0.985	0.805	0.876	0.846
PCA(2)-S	0.815	1.053	1.002	1.058	1.030	1.017	0.973	0.903	1.230	0.976	0.823	0.984	0.813	0.825	0.943
PCA(3)-S	0.388	1.052	1.004	1.107	1.014	1.102	0.964	1.018	1.376	0.828	1.241	1.052	0.768	0.878	1.198
PLS(1)-S	0.976	1.035	0.996	1.062	1.026	1.003	0.975	0.911	0.967	0.964	0.851	0.985	0.805	0.876	0.846
PLS(2)-S	0.815	1.053	1.002	1.058	1.030	1.017	0.973	0.903	1.230	0.976	0.823	0.984	0.813	0.825	0.943
PLS(3)-S	0.388	1.052	1.004	1.107	1.014	1.102	0.964	1.018	1.376	0.828	1.241	1.052	0.768	0.878	1.198
RF-S	0.769	1.021	1.035	1.262	0.990	1.106	0.996	0.896	0.976	0.901	0.935	0.970	0.931	0.966	0.884
MLP-S	1.027	1.114	1.028	0.982	1.223	1.043	0.956	0.724	0.991	1.681	0.579	1.045	1.301	0.625	0.808
EML-S	1.228	1.023	1.002	0.988	0.991	1.027	1.005	0.795	1.022	1.002	0.706	0.986	0.939	0.672	0.824
SVR-S	1.087	1.034	1.029	1.055	1.010	1.052	0.983	0.887	1.061	0.913	0.870	1.018	0.983	0.969	0.935
AVG-BFS-PEN-S	0.855	1.025	1.012	1.409	1.048	1.189	0.940	1.100	1.043	1.016	0.858	0.932	1.069	0.782	0.955
AVG-PCA-S	0.600	1.044	0.999	1.030	1.022	1.037	0.969	0.942	1.180	0.900	0.949	1.001	0.795	0.853	0.987
AVG-PLS-S	0.600	1.044	0.999	1.030	1.022	1.037	0.969	0.942	1.180	0.900	0.949	1.001	0.795	0.853	0.987
AVG-NL-S	0.804	1.031	1.023	1.068	1.052	1.054	0.985	0.820	1.007	0.667	0.767	1.004	1.031	0.779	0.861
AVG-ALL-S	0.634	1.035	1.006	1.105	1.035	1.070	0.965	0.931	1.081	0.769	0.878	0.983	0.911	0.812	0.943
AVG-ALL-wAR-S	0.944	1.024	1.002	1.053	1.025	1.035	0.972	0.916	1.020	0.742	0.906	0.997	0.975	0.899	1.057

Table 10: PSP nowcasting cross-validation results from 2016 to 2019 using MV1 pattern for missing values. Actual values are reported for AR(1). Relative values to the AR(1) benchmark are reported for all other models.

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E-AGG	POS1	POS2	POS3	POS-AGG
AR(1)	1.290	6.598	3.630	2.091	4.686	2.659	4.950	3.140	1.960	0.536	0.192	5.903	9.337	5.120	1.742
AR(P)	1.043	1.055	1.003	1.000	1.007	1.039	0.966	0.787	0.936	0.588	0.543	1.019	1.040	1.072	1.553
AVG-AR	1.114	1.030	1.002	1.000	1.005	1.026	0.957	0.833	0.956	0.607	0.522	1.009	1.020	1.048	1.367
BFS-L	2.199	1.362	1.375	0.964	0.876	1.349	0.856	1.251	1.233	3.315	3.026	1.089	1.477	1.060	1.246
Ridge-L	1.318	0.980	0.999	1.003	1.007	1.061	1.034	0.603	0.988	1.108	1.486	1.026	0.999	0.974	1.046
Lasso-L	1.166	0.997	1.001	1.002	1.022	1.161	0.832	0.892	0.944	2.483	2.073	1.031	1.047	0.864	0.931
EN05-L	1.076	1.003	1.031	1.066	1.006	1.185	0.824	0.856	1.029	1.425	1.506	1.022	1.025	0.854	0.899
AdRidge-L	0.852	0.945	1.009	0.853	0.971	1.302	0.837	0.917	1.102	2.120	2.241	1.195	1.159	0.706	0.917
AdLasso-L	0.784	0.957	1.023	0.907	1.011	1.264	0.837	0.898	0.976	3.202	2.373	0.968	1.132	0.705	0.893
AdEN05-L	1.216	0.945	1.049	1.005	0.998	1.272	0.825	0.922	0.939	2.416	2.346	0.994	1.047	0.676	0.872
PCA(1)-L	1.585	1.038	1.002	0.984	1.013	1.030	1.047	0.591	0.988	1.235	1.512	1.007	0.929	1.002	0.865
PCA(2)-L	1.590	1.035	0.988	1.005	1.007	1.055	1.066	0.597	0.971	1.239	1.667	1.016	0.933	1.017	0.832
PCA(3)-L	1.228	0.991	0.995	1.386	0.966	1.028	1.072	0.669	1.173	1.323	1.501	1.138	0.943	1.010	1.431
PLS(1)-L	1.585	1.038	1.002	0.984	1.013	1.030	1.047	0.591	0.988	1.235	1.512	1.007	0.929	1.002	0.865
PLS(2)-L	1.590	1.035	0.988	1.005	1.007	1.055	1.066	0.597	0.971	1.239	1.667	1.016	0.933	1.017	0.832
PLS(3)-L	1.228	0.991	0.995	1.386	0.966	1.028	1.072	0.669	1.173	1.323	1.501	1.138	0.943	1.010	1.431
RF-L	1.503	0.985	0.992	1.090	1.008	1.082	0.996	0.721	1.030	1.297	1.582	1.068	1.060	0.929	0.974
MLP-L	1.272	1.319	1.275	1.021	1.190	0.715	1.558	0.682	0.862	0.610	1.748	1.145	1.153	0.835	0.856
EML-L	1.388	1.052	1.078	1.013	0.980	1.091	1.065	0.602	0.995	1.116	2.316	0.998	0.939	0.954	0.880
SVR-L	1.386	1.007	0.991	1.005	1.011	1.086	1.036	0.665	1.051	0.912	1.466	1.037	0.987	1.079	0.949
AVG-BFS-PEN-L	1.070	1.014	1.053	0.875	0.943	1.190	0.855	0.860	0.979	2.160	2.017	1.036	1.106	0.777	0.924
AVG-PCA-L	1.433	1.019	0.995	1.118	0.994	1.037	1.062	0.616	1.040	1.262	1.416	1.049	0.935	1.009	1.037
AVG-PLS-L	1.433	1.019	0.995	1.118	0.994	1.037	1.062	0.616	1.040	1.262	1.416	1.049	0.935	1.009	1.037
AVG-NL-L	0.910	1.036	1.080	1.021	1.023	0.972	1.028	0.660	0.953	0.944	1.676	1.055	1.029	0.947	0.889
AVG-ALL-L	1.190	1.013	1.028	1.020	0.977	1.047	0.988	0.659	0.993	1.319	1.537	1.045	0.979	0.930	0.925
AVG-ALL-wAR-L	1.001	1.021	1.015	1.008	0.990	1.036	0.972	0.722	0.973	0.879	1.009	1.025	0.995	0.988	1.137
BFS-S	2.222	0.898	1.170	2.043	1.212	1.241	0.648	4.861	0.636	5.135	3.445	1.084	1.069	1.106	2.438
Ridge-S	1.335	1.044	1.003	1.000	1.007	1.040	1.046	0.603	1.016	1.107	1.653	1.025	0.934	0.955	1.062
Lasso-S	0.607	1.061	1.016	1.193	1.074	1.188	0.920	0.635	1.175	2.525	2.108	1.013	1.276	0.897	0.875
EN05-S	0.835	1.078	1.005	1.018	1.050	1.159	0.933	0.944	1.167	1.409	1.611	1.013	1.176	0.782	0.814
AdRidge-S	0.541	1.184	1.088	1.031	1.015	1.357	0.916	1.448	1.173	2.859	1.771	1.006	1.310	0.717	0.968
AdLasso-S	0.919	1.068	1.053	1.146	1.108	1.312	0.891	0.795	1.137	3.055	2.956	1.013	1.354	1.046	1.086
AdEN05-S	0.747	1.106	1.032	1.042	1.107	1.457	0.913	1.472	1.141	3.078	1.811	1.013	1.371	0.672	0.885
PCA(1)-S	1.778	1.057	1.003	0.987	1.019	1.038	1.063	0.598	0.956	1.151	1.480	1.010	0.923	1.041	0.871
PCA(2)-S	1.580	1.085	0.985	1.122	1.061	1.102	1.054	0.649	1.173	1.143	1.660	1.058	0.874	0.937	1.107
PCA(3)-S	1.679	1.031	1.015	1.301	1.018	1.073	1.058	0.842	1.197	1.339	2.412	1.052	0.843	1.022	1.288
PLS(1)-S	1.778	1.057	1.003	0.987	1.019	1.038	1.063	0.598	0.956	1.151	1.480	1.010	0.923	1.041	0.871
PLS(2)-S	1.580	1.085	0.985	1.122	1.061	1.102	1.054	0.649	1.173	1.143	1.660	1.058	0.874	0.937	1.107
PLS(3)-S	1.679	1.031	1.015	1.301	1.018	1.073	1.058	0.842	1.197	1.339	2.412	1.052	0.843	1.022	1.288
RF-S	0.809	1.023	1.038	1.175	1.024	1.135	1.050	0.697	0.982	1.295	1.725	1.018	0.935	0.981	0.906
MLP-S	1.227	1.320	1.054	1.019	1.312	0.718	1.524	0.697	0.836	0.621	2.659	1.156	1.043	0.771	0.865
EML-S	1.381	1.052	1.005	1.013	0.999	1.091	1.065	0.602	0.995	1.116	2.795	0.998	0.967	0.847	0.708
SVR-S	1.425	1.067	1.021	0.988	1.028	1.126	1.062	0.673	1.036	0.928	1.531	1.020	0.967	1.076	0.948
AVG-BFS-PEN-S	0.255	1.049	1.025	1.058	1.045	1.229	0.877	1.312	1.033	2.364	1.823	0.994	1.186	0.861	0.985
AVG-PCA-S	1.562	1.052	1.000	1.132	1.031	1.069	1.057	0.682	1.108	1.188	1.458	1.038	0.879	0.996	1.086
AVG-PLS-S	1.562	1.052	1.000	1.132	1.031	1.069	1.057	0.682	1.108	1.188	1.458	1.038	0.879	0.996	1.086
AVG-NL-S	0.744	1.056	1.024	1.031	1.045	0.993	1.035	0.661	0.938	0.930	2.018	1.040	0.975	0.909	0.744
AVG-ALL-S	0.979	1.047	1.010	1.044	1.029	1.080	0.991	0.740	1.030	1.269	1.622	1.025	0.962	0.930	0.916
AVG-ALL-wAR-S	0.897	1.039	1.006	1.017	1.016	1.052	0.974	0.739	0.989	0.828	1.053	1.015	0.987	0.987	1.138

Table 11: PSP nowcasting cross-validation results from 2016 to 2019 using MV2 pattern for missing values. Actual values are reported for AR(1). Relative values to the AR(1) benchmark are reported for all other models.

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E-AGG	POS1	POS2	POS3	POS-AGG
AR(1)	1.290	6.598	3.630	2.091	4.686	2.659	4.950	3.140	1.960	0.536	0.192	5.903	9.337	5.120	1.742
AR(P)	1.043	1.055	1.003	1.000	1.007	1.039	0.966	0.787	0.936	0.588	0.543	1.019	1.040	1.072	1.553
AVG-AR	1.114	1.030	1.002	1.000	1.005	1.026	0.957	0.833	0.956	0.607	0.522	1.009	1.020	1.048	1.367
BFS-L	2.311	0.880	1.295	1.529	0.982	0.848	0.813	0.632	1.287	3.781	2.698	1.451	1.663	0.817	1.107
Ridge-L	1.181	0.943	1.004	1.000	1.007	1.054	1.043	0.603	1.048	1.108	1.643	1.053	1.012	0.942	1.026
Lasso-L	1.320	0.979	1.095	1.021	0.967	1.196	0.867	0.522	1.050	1.820	1.860	1.024	1.141	0.821	0.807
EN05-L	1.207	0.977	1.057	1.013	0.995	1.116	0.888	0.534	1.115	1.418	2.058	1.015	1.046	0.986	0.765
AdRidge-L	1.507	0.952	1.103	1.301	1.166	1.288	0.897	0.507	1.301	2.337	2.430	1.050	1.032	0.842	0.715
AdLasso-L	1.342	0.935	1.219	1.161	0.875	1.243	0.840	0.488	1.087	2.757	2.073	1.016	1.183	0.707	0.645
AdEN05-L	1.086	0.936	1.100	1.240	0.916	1.284	0.795	0.544	1.190	2.303	2.585	1.084	1.036	0.911	0.645
PCA(1)-L	1.828	1.050	1.007	1.006	1.008	1.035	1.076	0.605	0.961	1.143	1.470	0.991	1.001	1.082	0.880
PCA(2)-L	1.530	1.018	0.974	0.975	1.014	1.060	1.082	0.630	1.002	1.249	1.951	1.004	0.881	0.973	0.644
PCA(3)-L	1.695	0.892	1.002	1.622	0.925	1.010	1.091	0.871	1.388	1.613	1.769	1.235	0.853	0.876	1.612
PLS(1)-L	1.828	1.050	1.007	1.006	1.008	1.035	1.076	0.605	0.961	1.143	1.470	0.991	1.001	1.082	0.880
PLS(2)-L	1.530	1.018	0.974	0.975	1.014	1.060	1.082	0.630	1.002	1.249	1.951	1.004	0.881	0.973	0.644
PLS(3)-L	1.695	0.892	1.002	1.622	0.925	1.010	1.091	0.871	1.388	1.613	1.769	1.235	0.853	0.876	1.612
RF-L	1.538	0.968	1.024	1.114	1.016	1.066	0.982	0.653	1.056	1.362	1.717	1.025	1.005	0.872	0.900
MLP-L	1.287	1.319	1.306	1.021	1.325	0.985	1.530	0.688	0.850	0.586	1.866	1.165	1.039	0.773	0.855
EML-L	1.395	1.052	1.130	1.013	1.027	1.049	1.065	0.602	0.995	1.116	2.345	0.998	0.967	0.850	0.880
SVR-L	1.643	1.016	0.998	1.046	1.033	1.113	1.048	0.644	1.048	1.025	1.319	1.031	1.014	1.128	0.971
AVG-BFS-PEN-L	1.259	0.935	1.114	1.144	0.951	1.134	0.842	0.534	1.106	2.120	2.091	1.091	1.128	0.843	0.673
AVG-PCA-L	1.325	0.979	0.994	1.187	0.975	1.033	1.082	0.679	1.105	1.300	1.410	1.064	0.902	0.973	1.031
AVG-PLS-L	1.325	0.979	0.994	1.187	0.975	1.033	1.082	0.679	1.105	1.300	1.410	1.064	0.902	0.973	1.031
AVG-NL-L	0.978	1.034	1.110	1.039	1.055	1.048	1.019	0.636	0.957	0.970	1.731	1.045	1.004	0.896	0.867
AVG-ALL-L	1.200	0.968	1.050	1.131	0.982	1.051	0.994	0.612	1.056	1.362	1.570	1.061	0.981	0.910	0.845
AVG-ALL-wAR-L	1.057	0.998	1.025	1.065	0.992	1.038	0.975	0.715	1.002	0.894	1.028	1.033	0.991	0.977	1.100
BFS-S	3.653	1.059	1.013	3.097	1.232	1.161	0.932	0.874	0.773	4.394	2.653	0.974	1.107	0.763	2.806
Ridge-S	1.192	1.038	1.003	1.000	1.007	1.044	1.072	0.603	1.047	1.113	1.584	1.031	0.911	0.928	0.940
Lasso-S	0.669	1.075	1.087	1.356	1.073	1.239	0.926	0.478	1.183	2.951	2.820	1.013	1.178	0.871	0.972
EN05-S	0.590	1.073	1.053	1.132	1.050	1.170	0.966	0.664	1.166	2.388	2.033	1.013	1.067	0.860	0.790
AdRidge-S	0.750	1.164	1.138	1.563	0.995	1.335	0.940	0.531	1.213	3.118	2.386	1.095	1.186	0.914	0.916
AdLasso-S	0.897	1.093	1.199	1.659	1.108	1.453	0.915	0.627	1.184	3.287	3.689	1.013	1.212	0.817	1.173
AdEN05-S	0.810	1.088	1.108	1.409	1.107	1.457	0.945	0.552	1.167	3.321	2.022	1.013	1.195	0.910	0.956
PCA(1)-S	1.574	1.062	1.004	0.968	1.045	1.037	1.040	0.615	0.946	1.257	1.561	0.993	0.864	0.979	0.668
PCA(2)-S	1.247	1.102	1.013	1.226	1.049	1.186	1.008	0.722	1.342	1.318	1.771	1.091	0.817	0.788	1.230
PCA(3)-S	1.796	1.031	1.043	1.421	0.990	1.120	1.041	0.999	1.327	1.680	3.190	1.064	0.816	0.926	1.377
PLS(1)-S	1.574	1.062	1.004	0.968	1.045	1.037	1.040	0.615	0.946	1.257	1.561	0.993	0.864	0.979	0.668
PLS(2)-S	1.247	1.102	1.013	1.226	1.049	1.186	1.008	0.722	1.342	1.318	1.771	1.091	0.817	0.788	1.230
PLS(3)-S	1.796	1.031	1.043	1.421	0.990	1.120	1.041	0.999	1.327	1.680	3.190	1.064	0.816	0.926	1.377
RF-S	1.035	1.035	1.030	1.264	1.016	1.102	1.014	0.722	0.980	1.493	1.630	1.026	0.878	0.873	0.842
MLP-S	1.245	1.312	1.050	1.019	1.284	0.713	1.513	0.712	1.196	0.656	2.290	1.051	1.064	0.773	1.011
EML-S	1.353	1.052	1.005	1.013	1.012	1.091	1.065	0.602	1.073	1.116	2.581	1.069	0.967	0.838	0.858
SVR-S	1.441	1.051	1.018	1.049	1.043	1.121	1.067	0.642	1.047	1.073	1.449	1.007	0.979	1.079	0.950
AVG-BFS-PEN-S	0.926	1.062	1.044	1.347	1.047	1.250	0.926	0.425	1.059	2.788	2.041	0.982	1.103	0.830	1.042
AVG-PCA-S	1.361	1.052	1.019	1.199	1.023	1.109	1.023	0.756	1.198	1.391	1.553	1.045	0.831	0.887	1.087
AVG-PLS-S	1.361	1.052	1.019	1.199	1.023	1.109	1.023	0.756	1.198	1.391	1.553	1.045	0.831	0.887	1.087
AVG-NL-S	0.806	1.054	1.022	1.076	1.052	0.985	1.030	0.657	1.057	1.021	1.892	1.025	0.971	0.881	0.773
AVG-ALL-S	0.847	1.048	1.013	1.172	1.027	1.106	0.987	0.583	1.113	1.540	1.692	1.021	0.917	0.849	0.933
AVG-ALL-wAR-S	0.891	1.039	1.007	1.084	1.015	1.065	0.972	0.703	1.031	0.942	1.085	1.013	0.960	0.947	1.143

Table 12: PSP nowcasting cross-validation results from 2016 to 2019 using MV3 pattern for missing values. Actual values are reported for AR(1). Relative values to the AR(1) benchmark are reported for all other models.

	2016	**2017**	**2018**	**2019**
gbdemo1180	-0.348	-0.35	-0.329	-0.308
gb_swiid_rel_red	0.245	0.208	0.292	0.292
uspoli0211	-0.474	-0.485	-0.416	-0.416
uspoli0583	-0.156	-0.291	-0.213	-0.137
uspoli1513	-0.081	NA	-0.066	-0.065
uspoli0769	NA	-0.093	-0.011	NA
gbfofi0103	0.301	0.175	0.283	0.287
gblama1207	0.097	0.206	0.106	0.108
gbmost0100	0.766	NA	0.694	0.655
gbmost0046	NA	0.911	NA	NA
gbprod0515	-0.024	-0.053	-0.013	NA
gbcaes0058	0.145	0.073	0.138	0.138

Table 13: EN05-S beta coefficient estimates using MV3 missing values pattern for E1 target.

	2016	2017	2018	2019
gbdemo0236	NA	NA	NA	1.043
uspoli1699	NA	0.217	NA	NA
gb_pv_est_esg	-0.01	-0.005	NA	-0.011
gbbank02082	0.047	NA	NA	NA
gbnaac00352	0.125	NA	NA	0.161
gbnaac00422	NA	NA	-0.062	NA
ukepunnewsindex	NA	0.016	0.018	0.018
gbprod0515	NA	NA	-0.108	NA
gbsurv0559	NA	-0.058	NA	NA
csiigbp	0.054	NA	NA	NA
lse_all_mcap	0.221	0.14	0.166	0.192
ecb_cissdgbz0z4fecss_cinidx	NA	NA	2.256	NA

Table 14: BFS-S beta coefficient estimates using MV3 missing values pattern for E9 target.

	2016	**2017**	**2018**	**2019**
bp_gbgasconscm	NA	NA	NA	-0.232
bp_gbenergyconsej	-0.66	NA	NA	-0.5
gbdemo0233	NA	0.69	NA	2.685
gbdemo0230	NA	-1.215	NA	NA
gb_swiid_gini_disp	-0.468	NA	NA	NA
gb_rl_est_esg	0.297	NA	NA	NA
gbwui	0.062	NA	NA	NA
gbnaac00332	NA	-2.138	NA	NA
gbnaac00342	NA	NA	NA	0.661
gbcaes0167	NA	NA	-3.194	NA
gbmost0100	NA	NA	0.174	NA
gbbank6836	NA	NA	0.232	0.246
gbpric00211	-1.188	-3.286	NA	NA
gbsurv0520	NA	NA	-0.064	NA
gbpeur	NA	NA	-0.575	NA
gbp	NA	-0.285	NA	NA

Table 15: BFS-S beta coefficient estimates using MV3 missing values pattern for POS1 target.

Target	Top Model (MV3)	2020	2021	2022
E1	EN05-S	164.644	169.440	171.946
E2	BFS-L	147.735	149.170	152.626
E3	PCA(2)-L	131.503	132.671	133.819
E4	PCA(1)-S	176.845	180.614	188.446
E5	AdLasso-L	182.245	184.865	187.234
E6	MLP-S	180.417	183.442	186.366
E7	AdEN05-L	138.490	138.890	138.086
E8	AVG-BFS-PEN-S	159.066	161.475	165.590
E9	BFS-S	93.969	95.383	92.431
E10	MLP-L	103.880	103.937	103.966
E-AGG	AVG-AR	139.689	140.666	141.547
POS1	BFS-S	96.447	94.396	87.320
POS2	PCA(3)-S	46.464	61.084	73.937
POS3	AdLasso-L	201.682	204.643	206.466
POS-AGG	PCA(2)-L	78.360	75.407	75.797

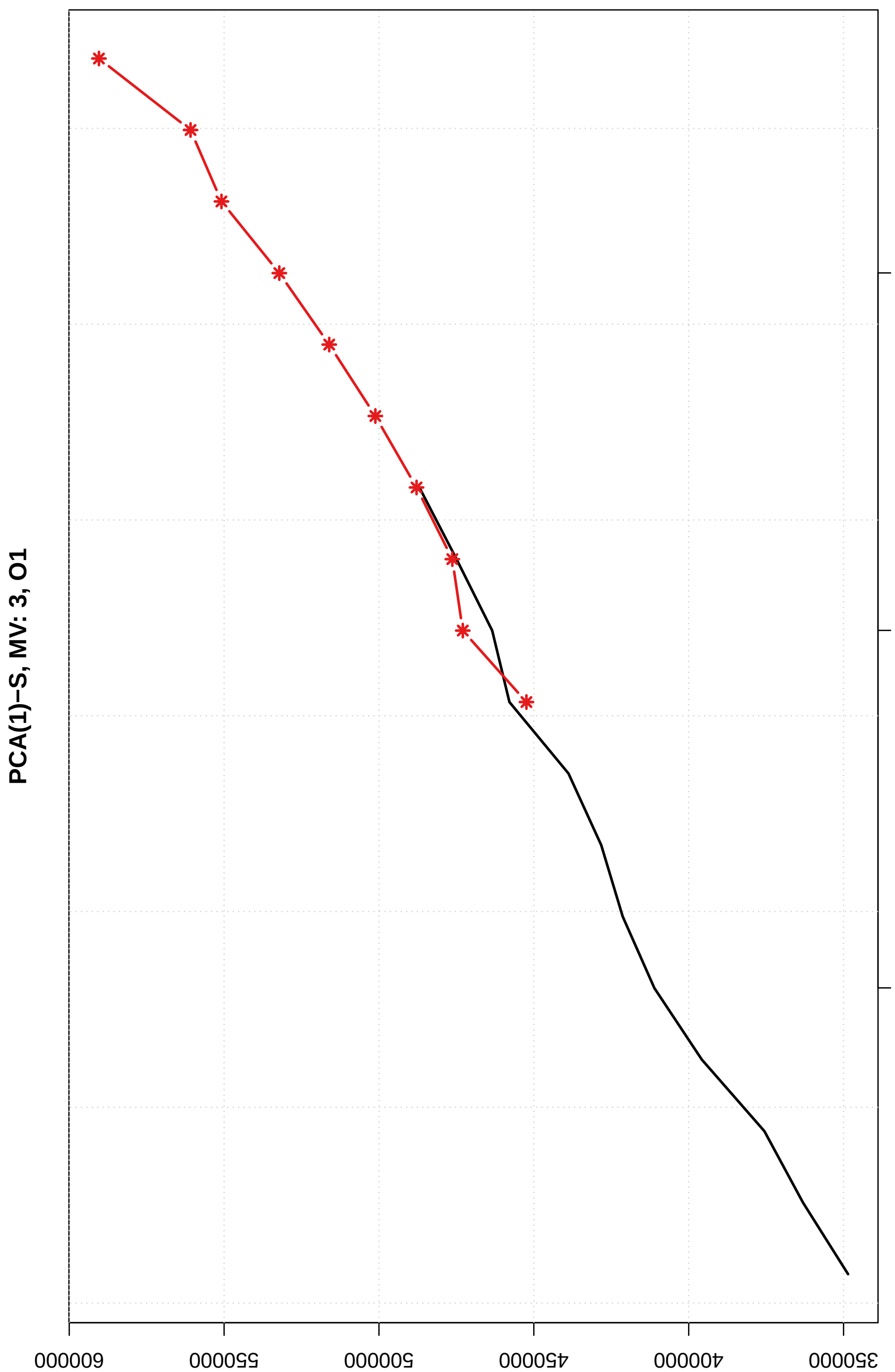
Table 16: PSP out-of-sample estimates for the top model of MV3 cross-validation. Illustrations for the above models are provided in the Appendix.

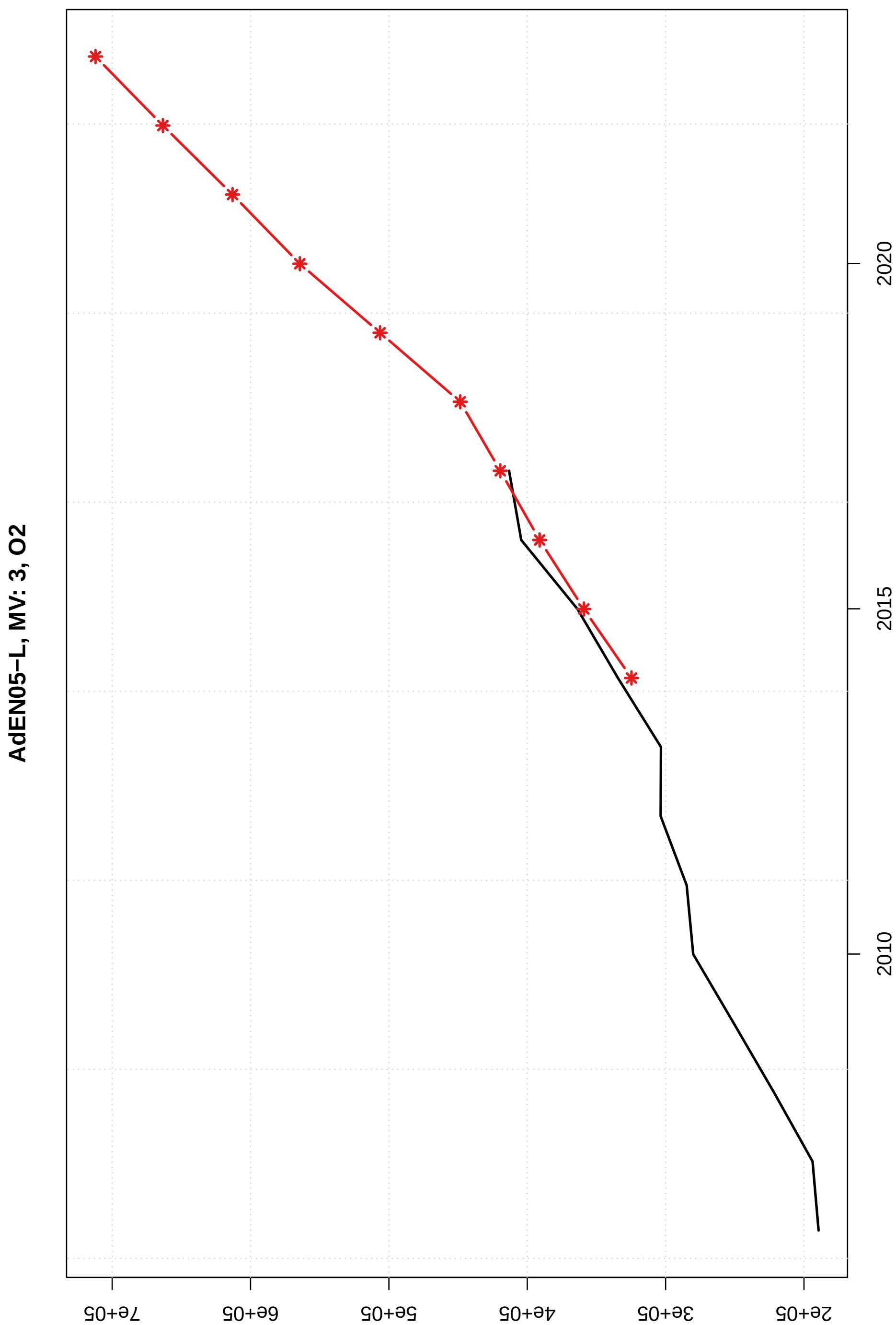
	MV1				MV2				MV3			
	E	**ASC**	**CSC**	**POS**	**E**	**ASC**	**CSC**	**POS**	**E**	**ASC**	**CSC**	**POS**
AR(1)	1.584	2.464	1.178	4.817	0.192	1.266	2.111	1.742	0.192	1.266	2.111	1.742
AR(P)	0.905	0.939	3.217	1.266	0.543	1.059	0.971	1.553	0.543	1.059	0.971	1.553
AVG-AR	0.936	0.959	2.199	1.176	0.522	1.039	0.870	1.367	0.522	1.039	0.870	1.367
BFS-L	1.349	0.927	5.771	1.469	3.026	1.299	4.446	1.246	2.698	1.479	5.386	1.107
Ridge-L	0.875	0.983	1.761	1.004	1.486	1.028	1.267	1.046	1.643	1.062	0.944	1.026
Lasso-L	0.791	0.993	2.403	0.985	2.073	1.074	1.999	0.931	1.860	1.088	1.568	0.807
EN05-L	0.803	0.986	2.676	0.905	1.506	1.050	1.964	0.899	2.058	1.050	0.951	0.765
AdRidge-L	0.840	0.960	3.129	0.857	2.241	1.000	2.507	0.917	2.430	0.981	1.569	0.715
AdLasso-L	0.801	0.969	3.154	0.979	2.373	1.064	2.174	0.893	2.073	1.072	2.629	0.645
AdEN05-L	0.764	0.969	2.908	0.861	2.346	1.074	2.468	0.872	2.585	1.088	1.911	0.645
PCA(1)-L	0.895	0.984	2.486	0.940	1.512	1.063	1.387	0.865	1.470	1.033	1.136	0.880
PCA(2)-L	0.896	1.020	1.944	0.843	1.667	1.061	1.418	0.832	1.951	1.120	1.629	0.644
PCA(3)-L	1.062	0.886	2.160	1.119	1.501	0.795	1.743	1.431	1.769	0.770	2.339	1.612
PLS(1)-L	0.895	0.984	2.486	0.940	1.512	1.063	1.387	0.865	1.470	1.033	1.136	0.880
PLS(2)-L	0.896	1.020	1.944	0.843	1.667	1.061	1.418	0.832	1.951	1.120	1.629	0.644
PLS(3)-L	1.062	0.886	2.160	1.119	1.501	0.795	1.743	1.431	1.769	0.770	2.339	1.612
RF-L	0.926	0.992	1.718	0.942	1.582	0.972	1.075	0.974	1.717	0.926	1.223	0.900
MLP-L	0.653	1.037	1.960	0.841	1.748	1.525	1.477	0.856	1.866	1.539	1.277	0.855
EML-L	0.778	1.068	1.941	0.901	2.316	1.153	0.892	0.880	2.345	1.153	0.892	0.880
SVR-L	0.889	1.037	1.893	0.931	1.466	1.053	0.735	0.949	1.319	1.056	0.570	0.971
AVG-BFS-PEN-L	0.880	0.967	2.850	0.997	2.017	1.076	2.155	0.924	2.091	1.090	1.988	0.673
AVG-PCA-L	0.950	0.963	1.956	0.965	1.416	0.966	1.505	1.037	1.410	0.919	1.662	1.031
AVG-PLS-L	0.950	0.963	1.956	0.965	1.416	0.966	1.505	1.037	1.410	0.919	1.662	1.031
AVG-NL-L	0.804	1.033	1.136	0.901	1.676	1.152	0.753	0.889	1.731	1.144	0.812	0.867
AVG-ALL-L	0.893	0.981	1.908	0.953	1.537	1.034	1.391	0.925	1.570	1.007	1.349	0.845
AVG-ALL-wAR-L	0.914	0.970	1.713	1.062	1.009	1.036	1.047	1.137	1.028	1.022	1.015	1.100
BFS-S	1.188	0.992	5.413	1.631	3.445	1.541	3.530	2.438	2.653	2.033	4.688	2.806
Ridge-S	0.853	0.980	1.658	0.976	1.653	1.006	1.274	1.062	1.584	0.988	1.141	0.940
Lasso-S	0.801	0.977	3.178	0.842	2.108	1.002	2.240	0.875	2.820	1.000	2.424	0.972
EN05-S	0.888	0.982	3.114	0.850	1.611	1.011	2.272	0.814	2.033	1.021	2.090	0.790
AdRidge-S	0.804	0.984	3.227	0.813	1.771	0.857	2.248	0.968	2.386	0.795	2.461	0.916
AdLasso-S	0.705	0.963	3.357	0.814	2.956	0.939	2.180	1.086	3.689	0.932	2.813	1.173
AdEN05-S	0.822	0.963	3.388	0.847	1.811	0.939	2.214	0.885	2.022	0.932	2.656	0.956
PCA(1)-S	0.851	1.054	1.695	0.846	1.480	1.053	1.177	0.871	1.561	1.046	1.637	0.668
PCA(2)-S	0.823	1.084	1.668	0.943	1.660	0.992	1.259	1.107	1.771	0.915	1.510	1.230
PCA(3)-S	1.241	0.851	4.944	1.198	2.412	0.732	1.879	1.288	3.190	0.850	2.007	1.377
PLS(1)-S	0.851	1.054	1.695	0.846	1.480	1.053	1.177	0.871	1.561	1.046	1.637	0.668
PLS(2)-S	0.823	1.084	1.668	0.943	1.660	0.992	1.259	1.107	1.771	0.915	1.510	1.230
PLS(3)-S	1.241	0.851	4.944	1.198	2.412	0.732	1.879	1.288	3.190	0.850	2.007	1.377
RF-S	0.935	1.018	1.977	0.884	1.725	0.970	1.196	0.906	1.630	1.004	1.360	0.842
MLP-S	0.579	1.044	2.704	0.808	2.659	1.706	1.362	0.865	2.290	1.821	1.385	1.011
EML-S	0.706	1.068	1.941	0.824	2.795	1.258	0.892	0.708	2.581	1.316	0.892	0.858
SVR-S	0.870	1.032	1.833	0.935	1.531	1.036	0.768	0.948	1.449	1.017	0.673	0.950
AVG-BFS-PEN-S	0.858	0.954	2.861	0.955	1.823	0.852	2.195	0.985	2.041	0.894	2.535	1.042
AVG-PCA-S	0.949	0.994	2.226	0.987	1.458	0.882	1.414	1.086	1.553	0.836	1.709	1.087
AVG-PLS-S	0.949	0.994	2.226	0.987	1.458	0.882	1.414	1.086	1.553	0.836	1.709	1.087
AVG-NL-S	0.767	1.040	1.272	0.861	2.018	1.198	0.934	0.744	1.892	1.231	0.903	0.773
AVG-ALL-S	0.878	0.993	1.806	0.943	1.622	0.933	1.399	0.916	1.692	0.913	1.483	0.933
AVG-ALL-wAR-S	0.906	0.976	1.312	1.057	1.053	0.984	1.049	1.138	1.085	0.972	1.088	1.143

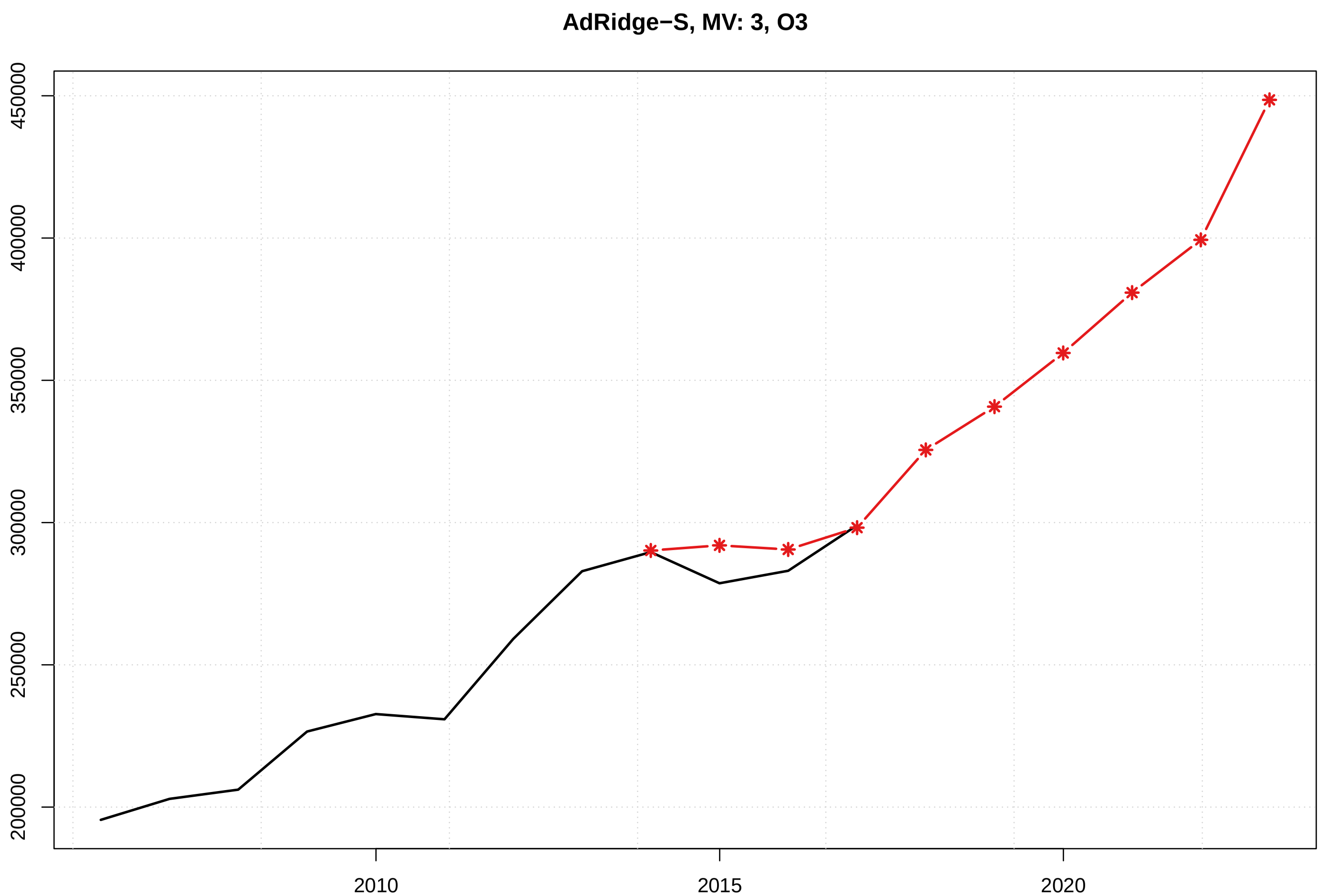
Table 17: PSP nowcasting cross-validation results from 2016 to 2019 for the final quality adjusted target using the three patterns for missing values. Actual values are reported for AR(1). Relative values to the AR(1) benchmark are reported for all other models.

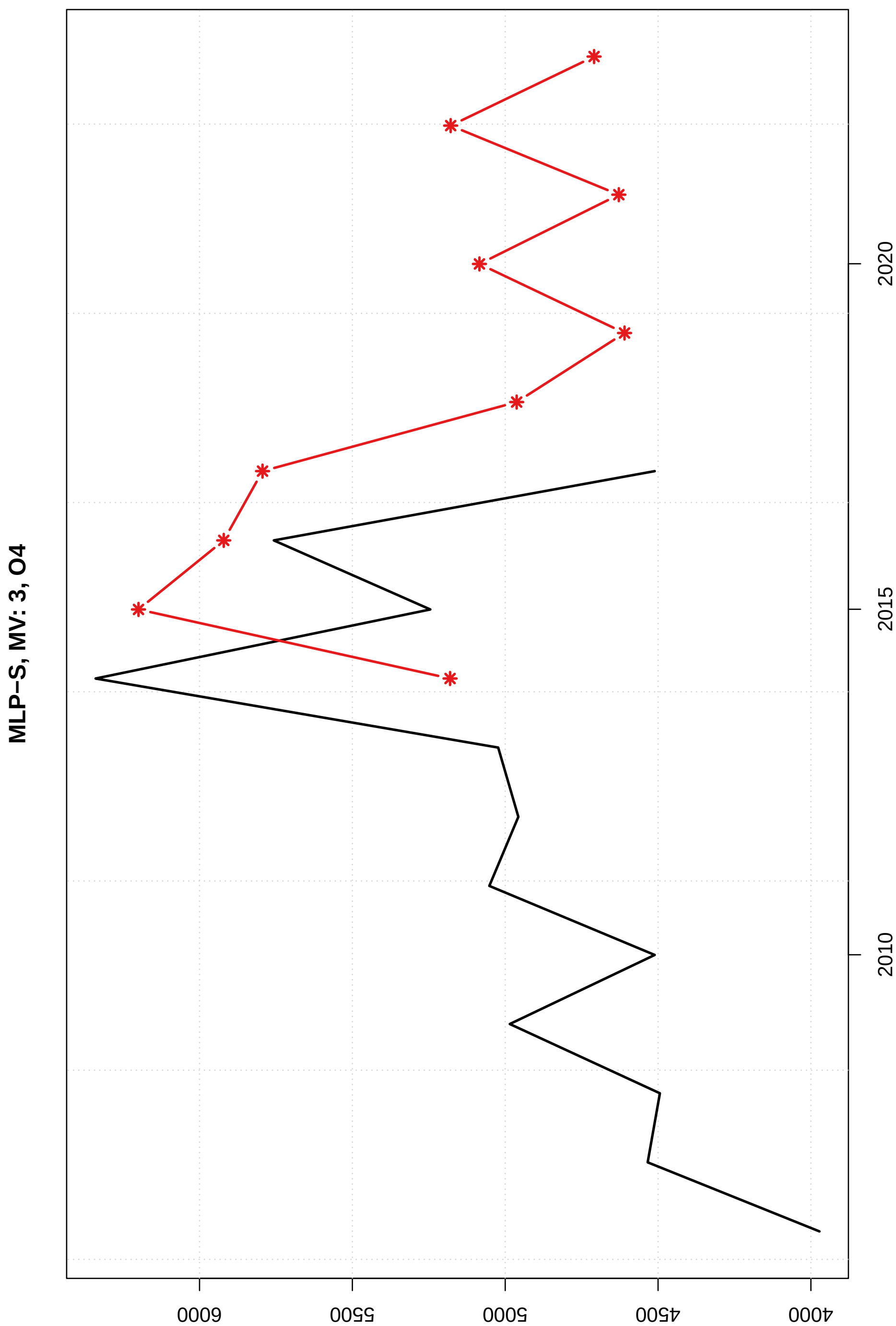
Appendix: Illustrative examples

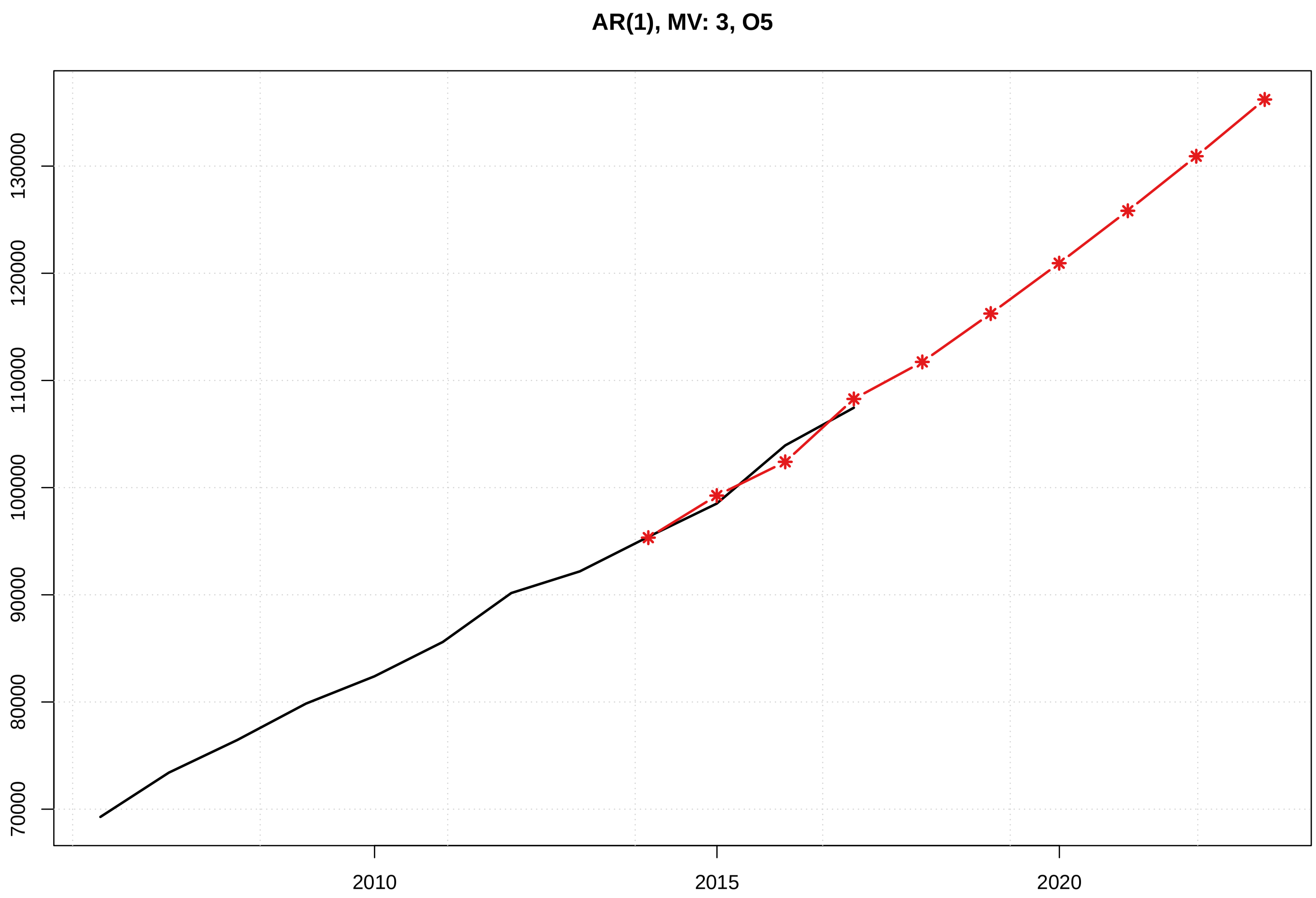
The figures which follow present the cross-validation (overlapping period) and real out-of-sample nowcasts for the best model for each UHP and PSP target and complements the information shown in Tables 9 and 16.

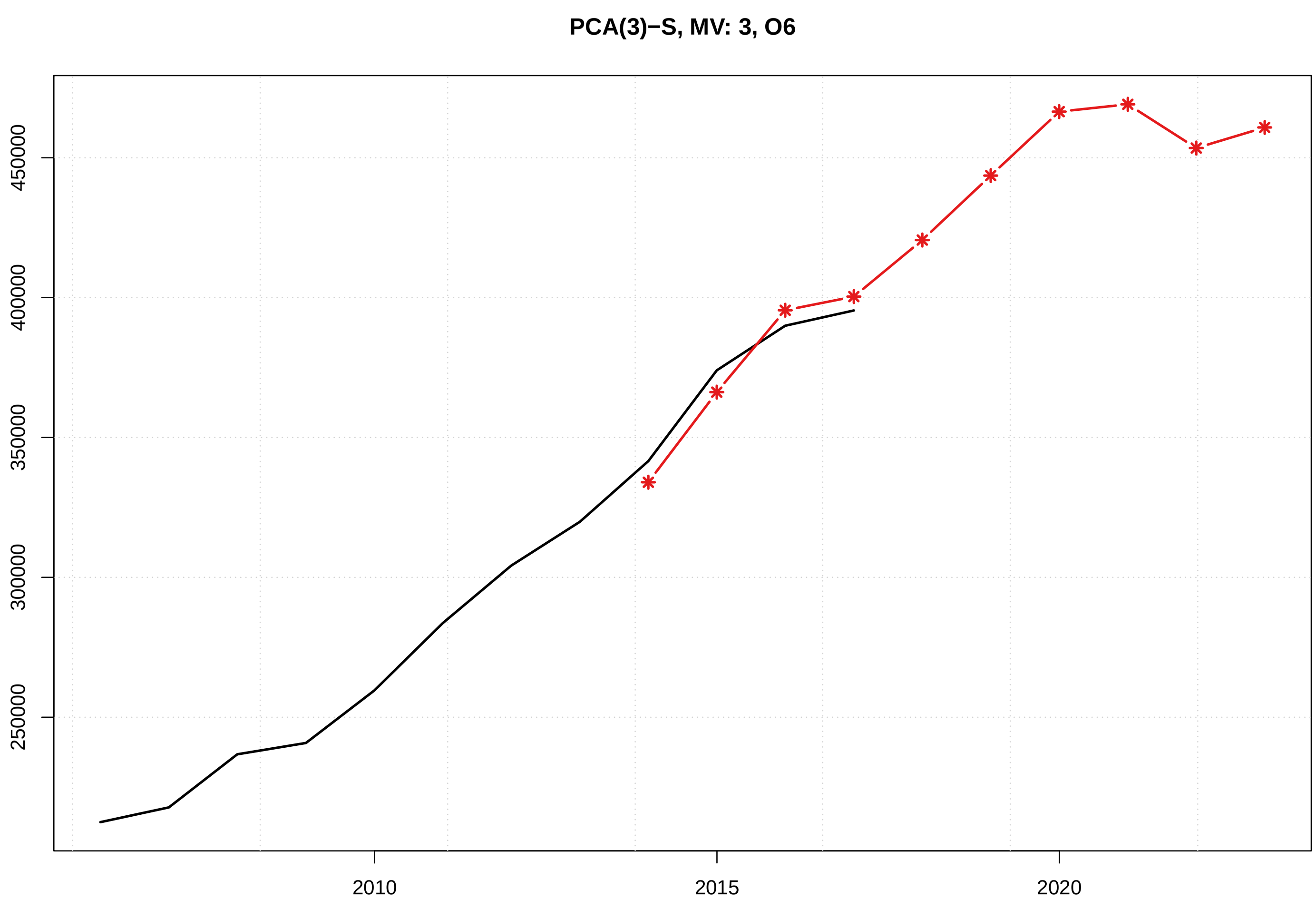

PCA(1)−S, MV: 3, O1
350000
400000
450000
500000
550000
600000
2010
2015
2020

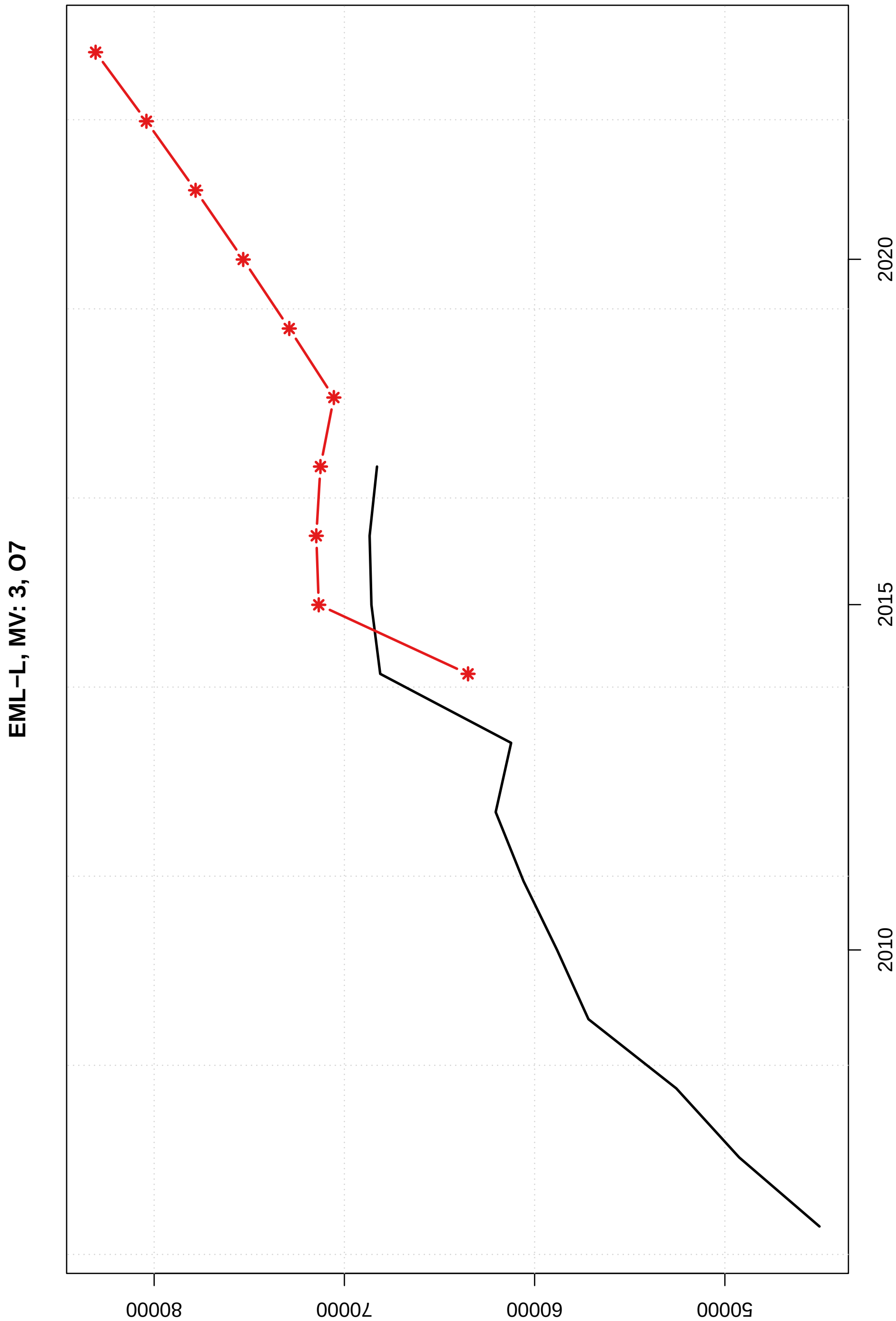

AdEN05−L, MV: 3, O2
2010
2015
2020
2e+05
3e+05
4e+05
5e+05
6e+05
7e+05


AdRidge−S, MV: 3, O3
450000
400000
350000
300000
250000
200000
2010
2015
2020

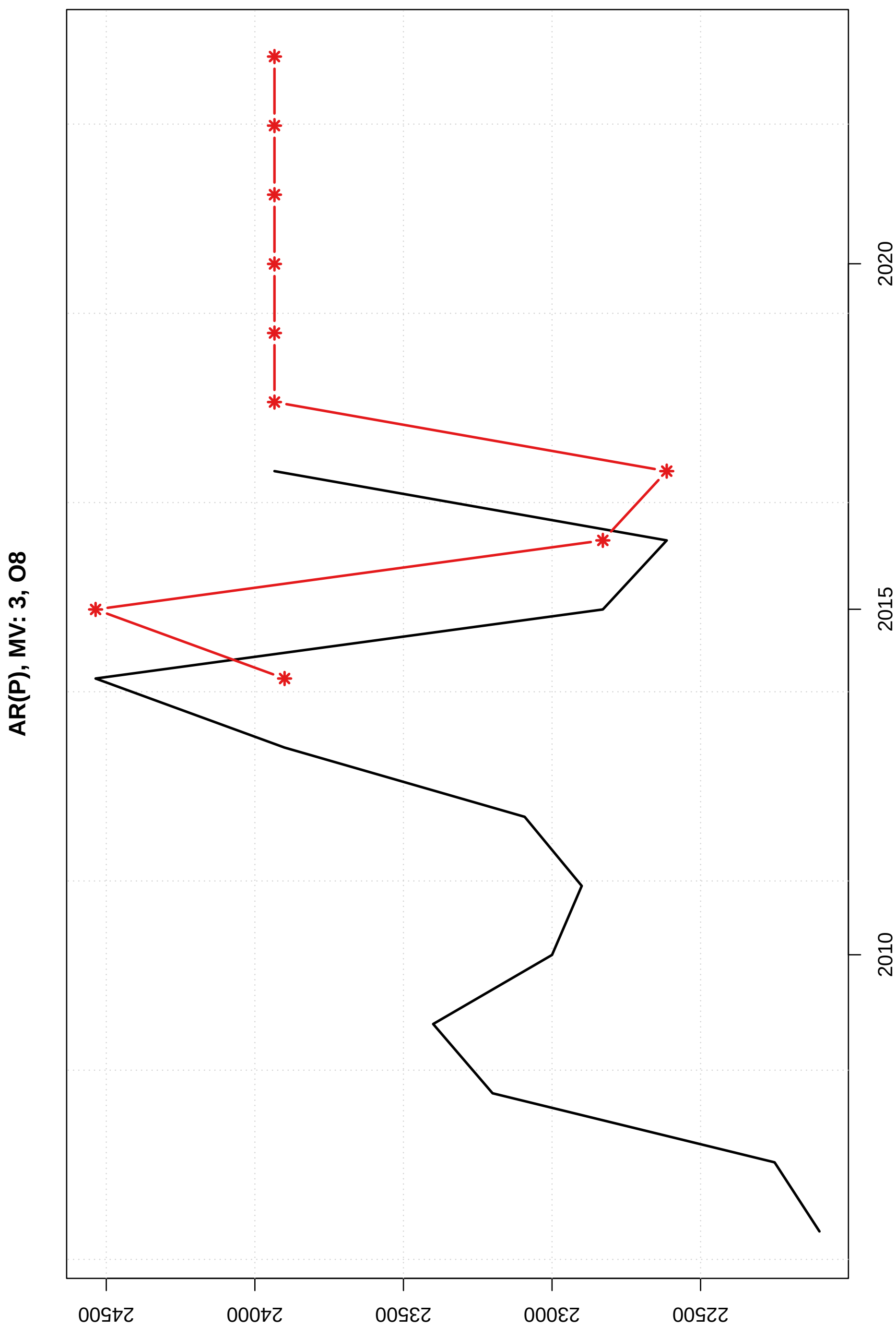

MLP−S, MV: 3, O4
4000
4500
5000
5500
6000
2010
2015
2020

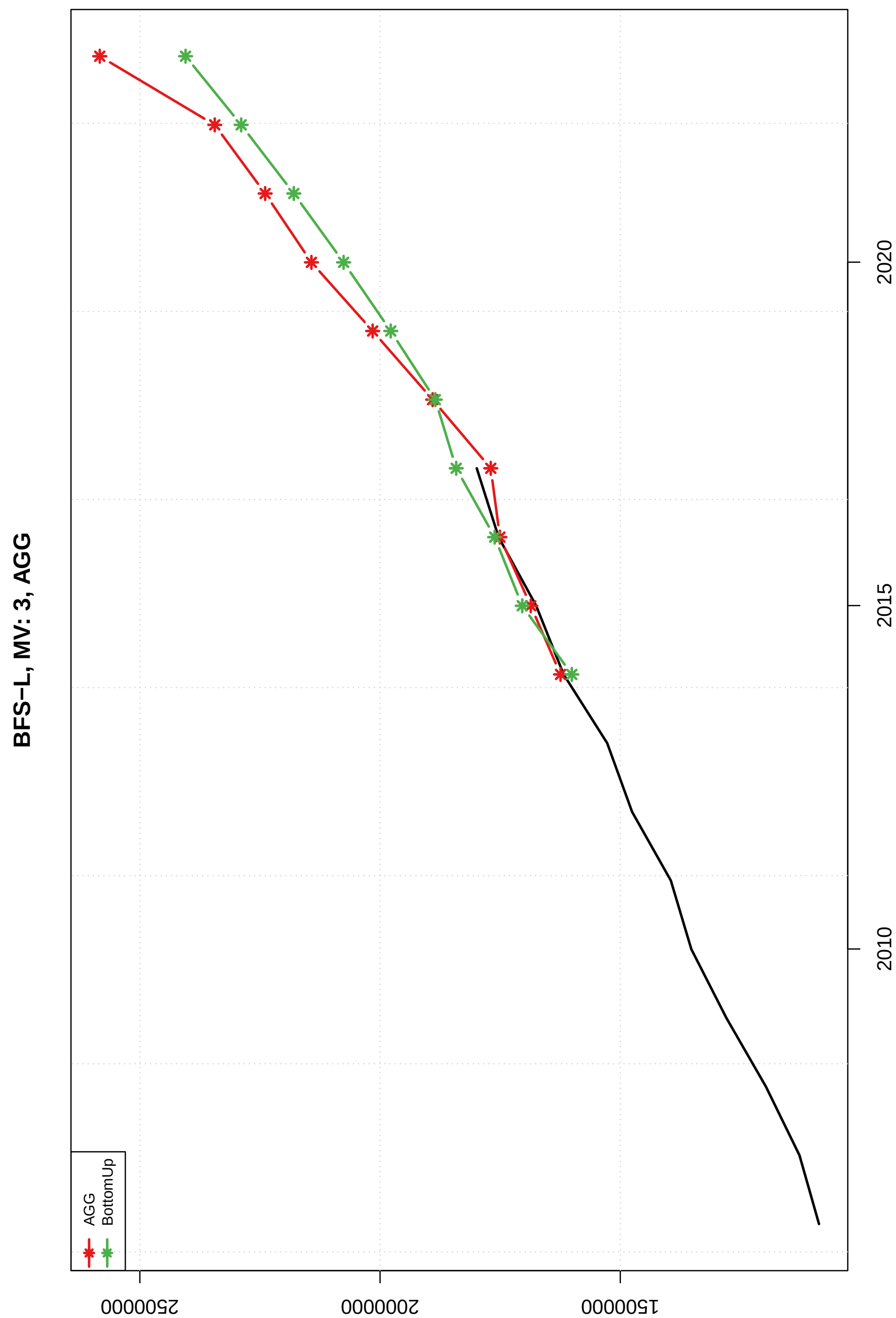

AR(1), MV: 3, O5
70000
80000
90000
100000
110000
120000
130000
2010
2015
2020

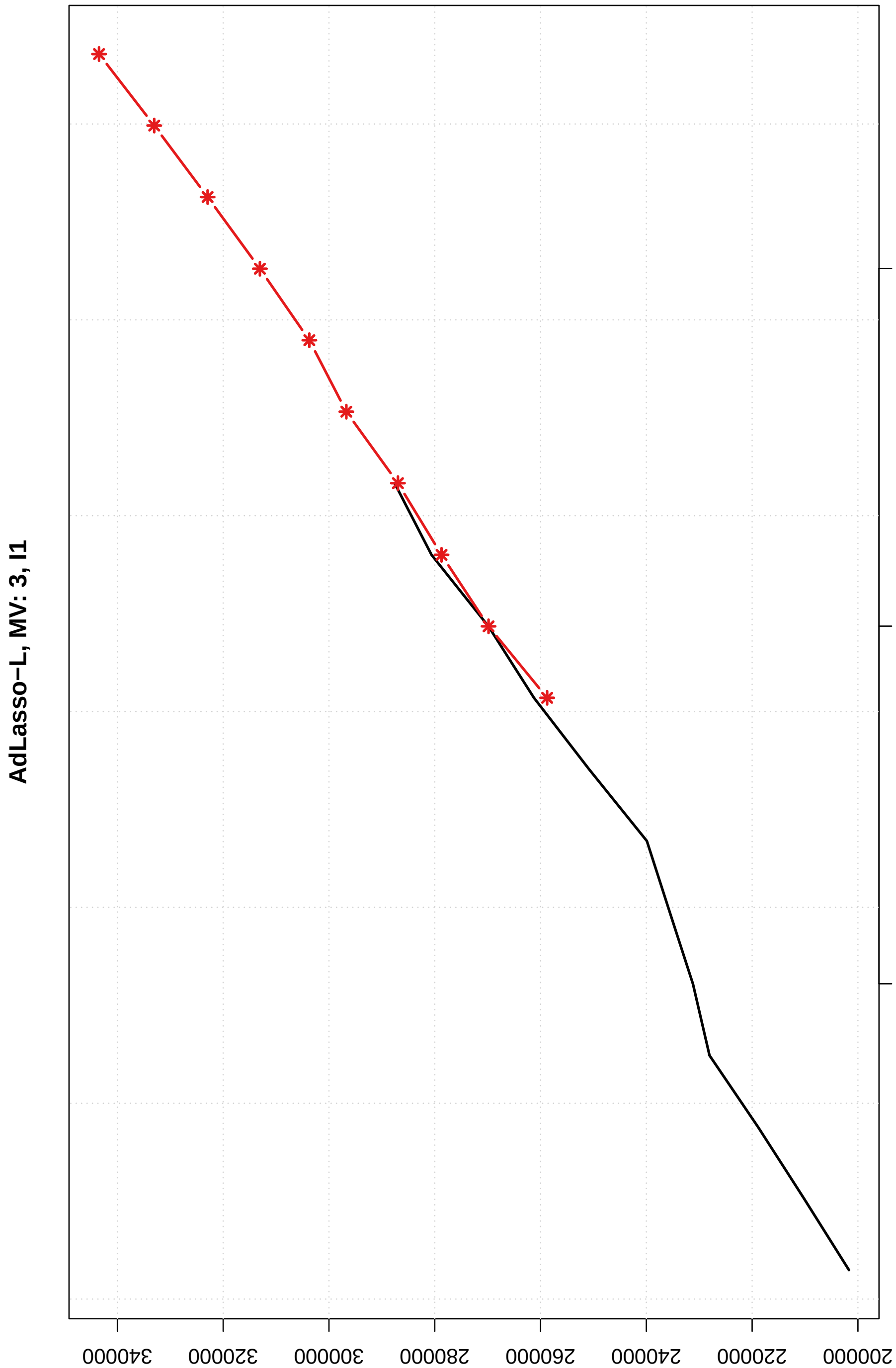

PCA(3)−S, MV: 3, O6
450000
400000
350000
300000
250000
2010
2015
2020

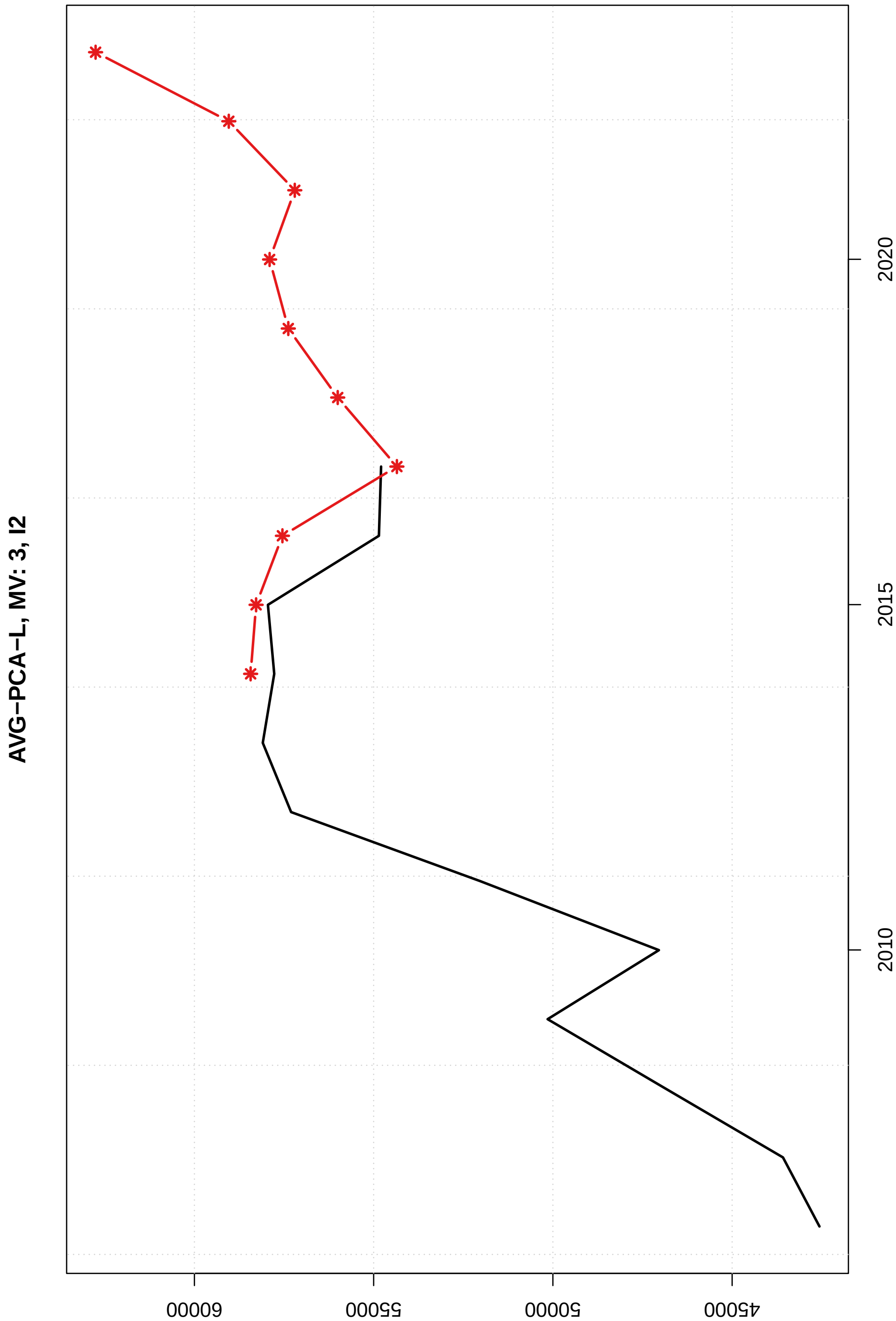


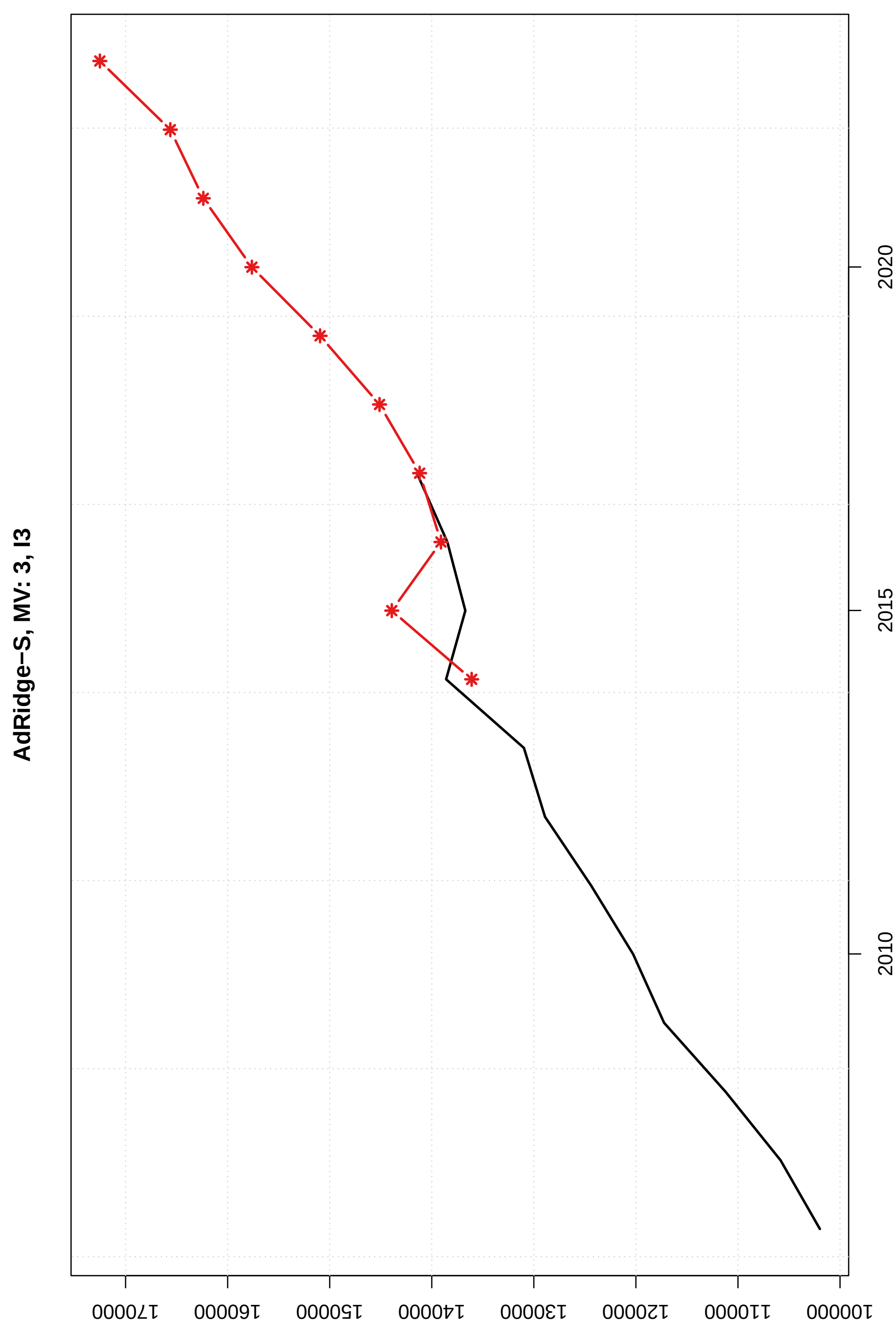
EML−L, MV: 3, O7
80000
70000
60000
50000
2010
2015
2020

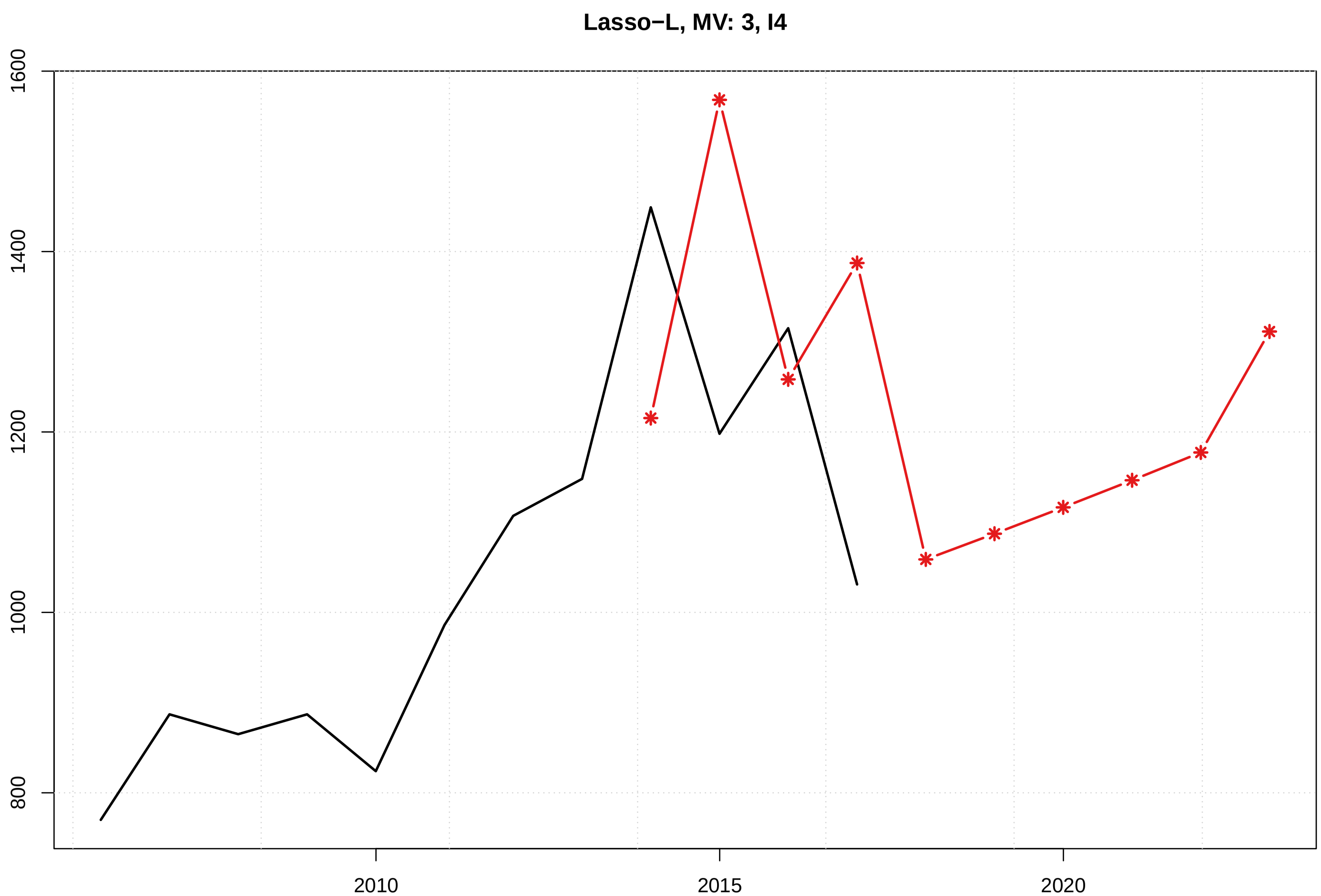

AR(P), MV: 3, O8
22500
23000
23500
24000
24500
2010
2015
2020

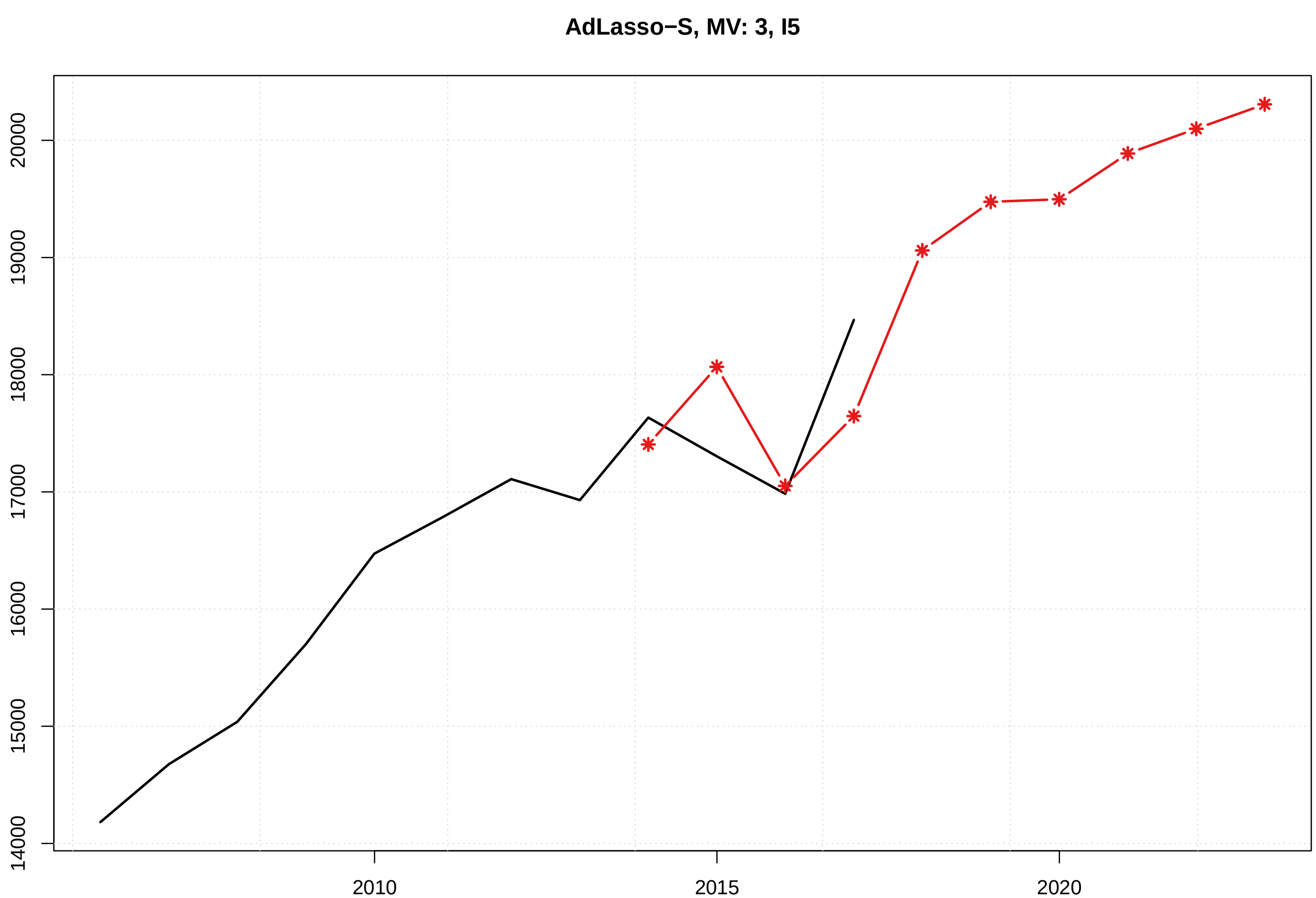

BFS−L, MV: 3, AGG
AGG
BottomUp
1500000
2000000
2500000
2010
2015
2020

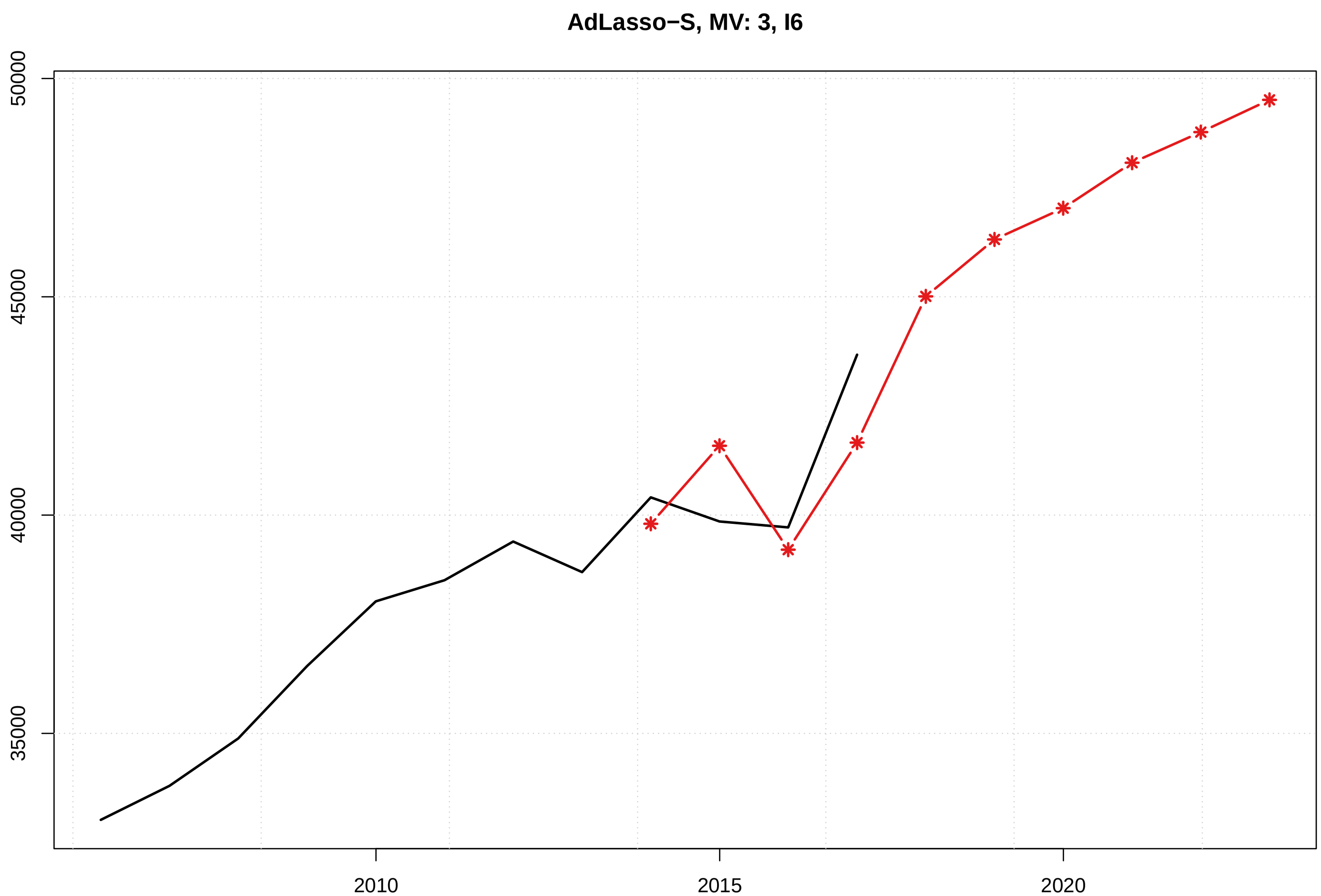

AdLasso−L, MV: 3, I1
200000
220000
240000
260000
280000
300000
320000
340000
2010
2015
2020


AVG−PCA−L, MV: 3, l2
45000
50000
55000
60000
2010
2015
2020

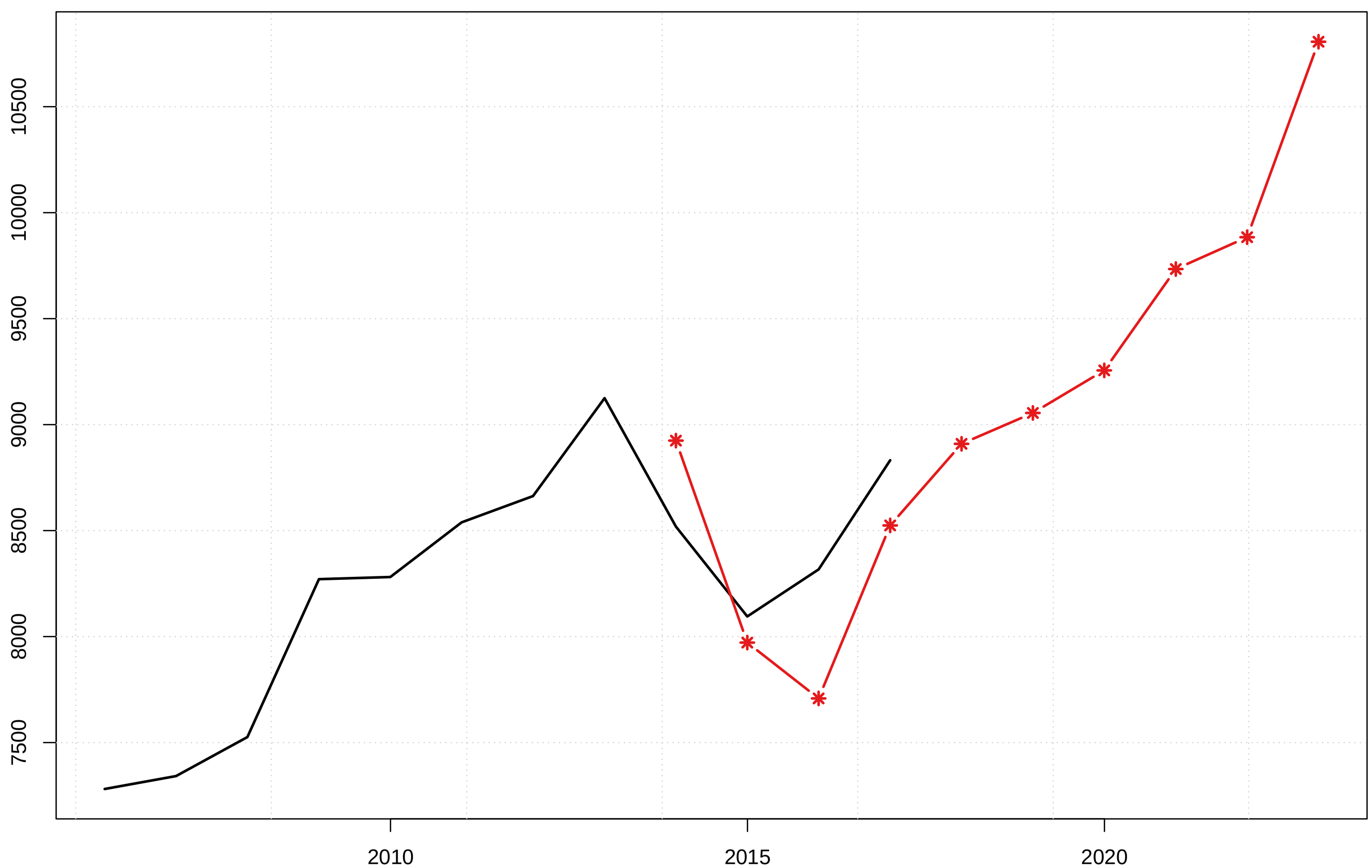

AdRidge−S, MV: 3, I3
2010
2015
2020
100000
110000
120000
130000
140000
150000
160000
170000

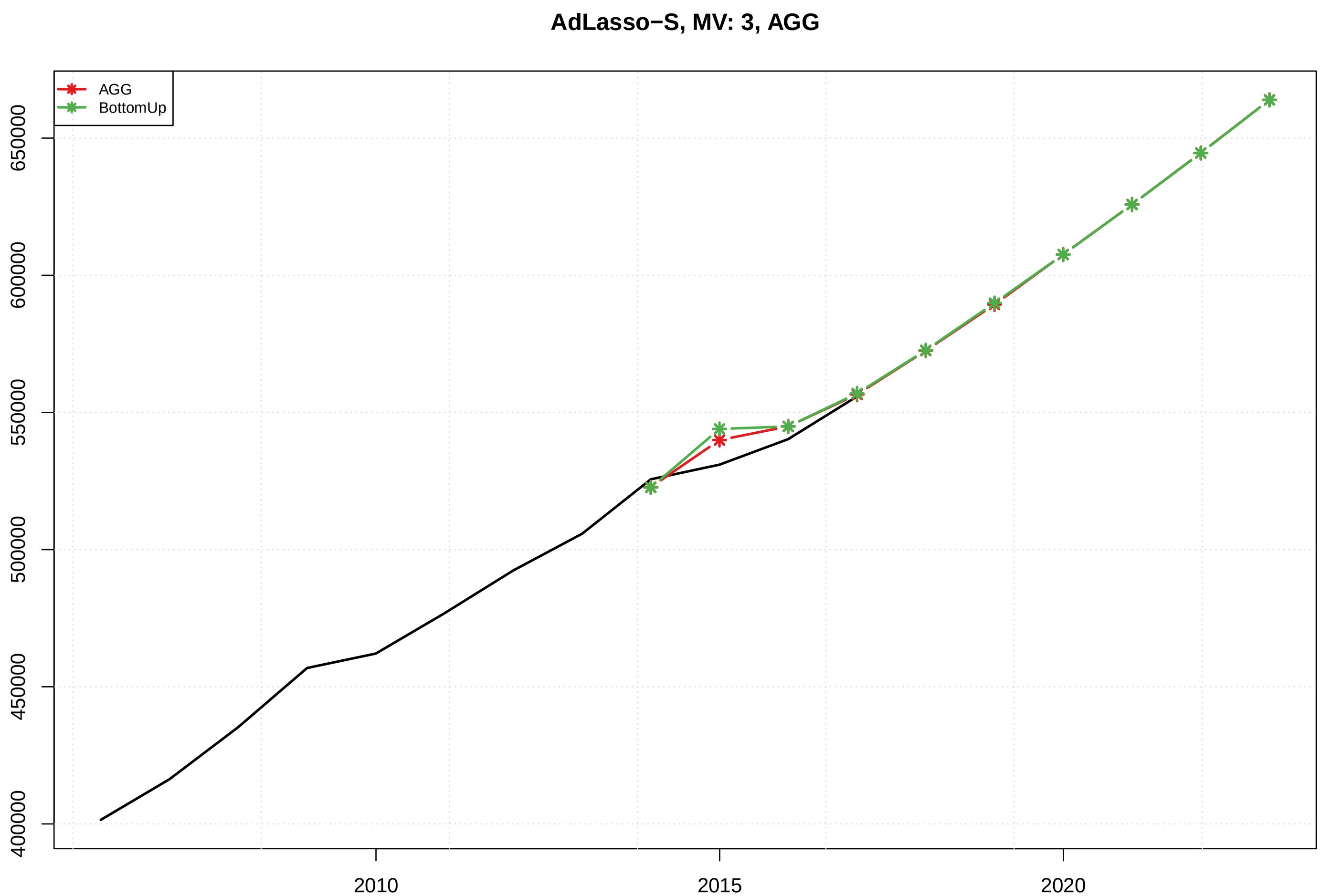

Lasso−L, MV: 3, I4
1600
1400
1200
1000
800
2010
2015
2020

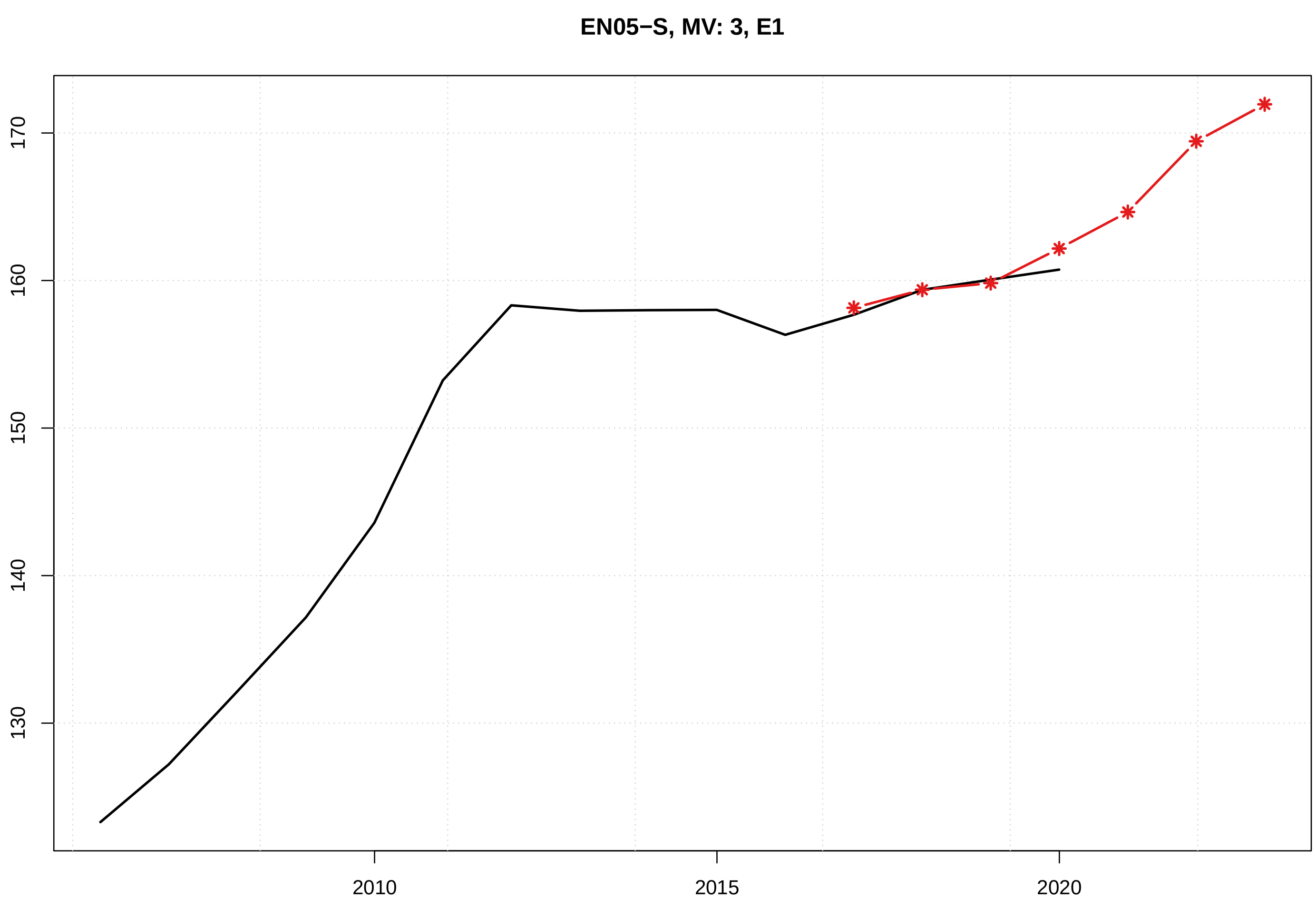

AdLasso−S, MV: 3, I5
14000
15000
16000
17000
18000
19000
20000
2010
2015
2020

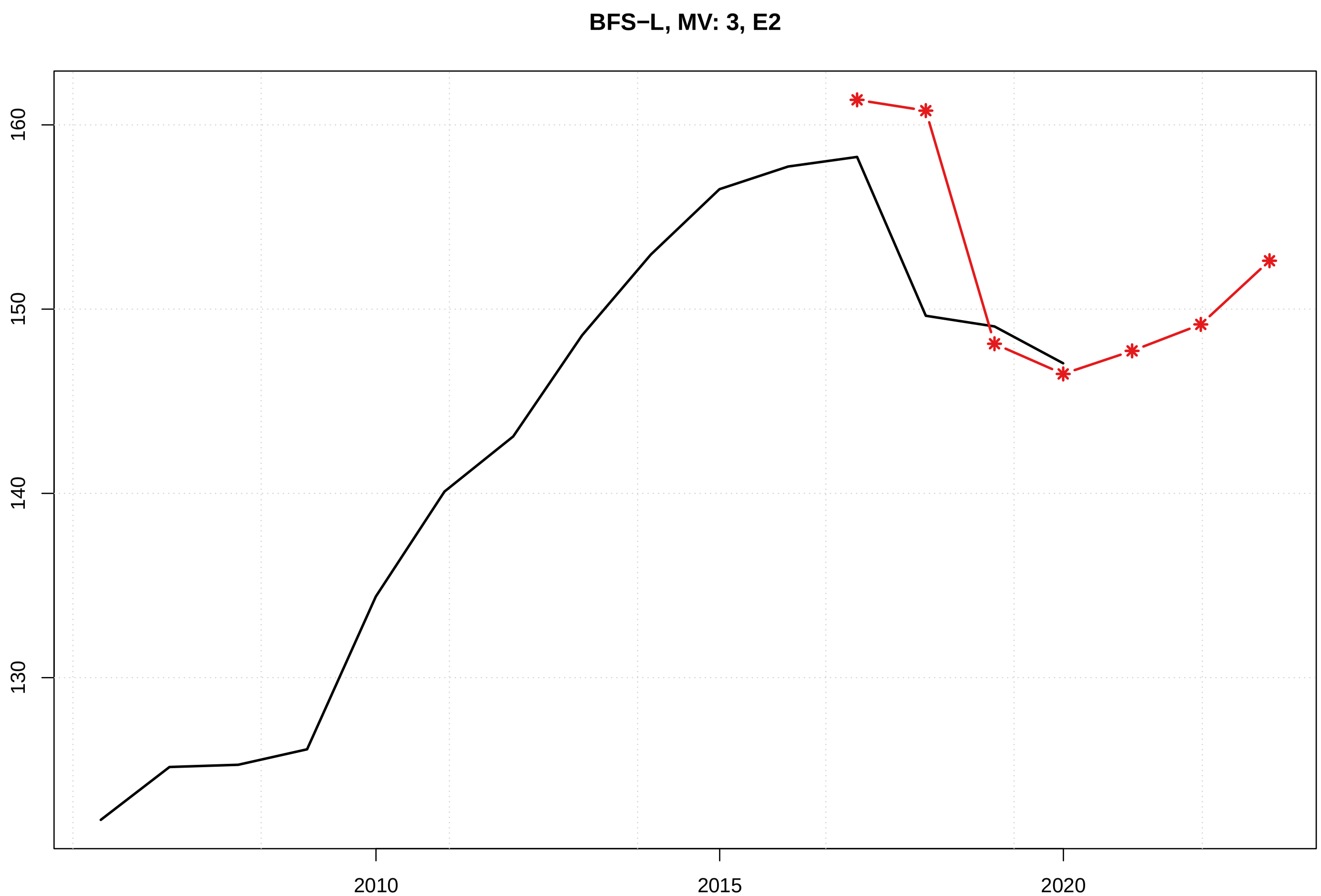

AdLasso−S, MV: 3, I6
50000
45000
40000
35000
2010
2015
2020

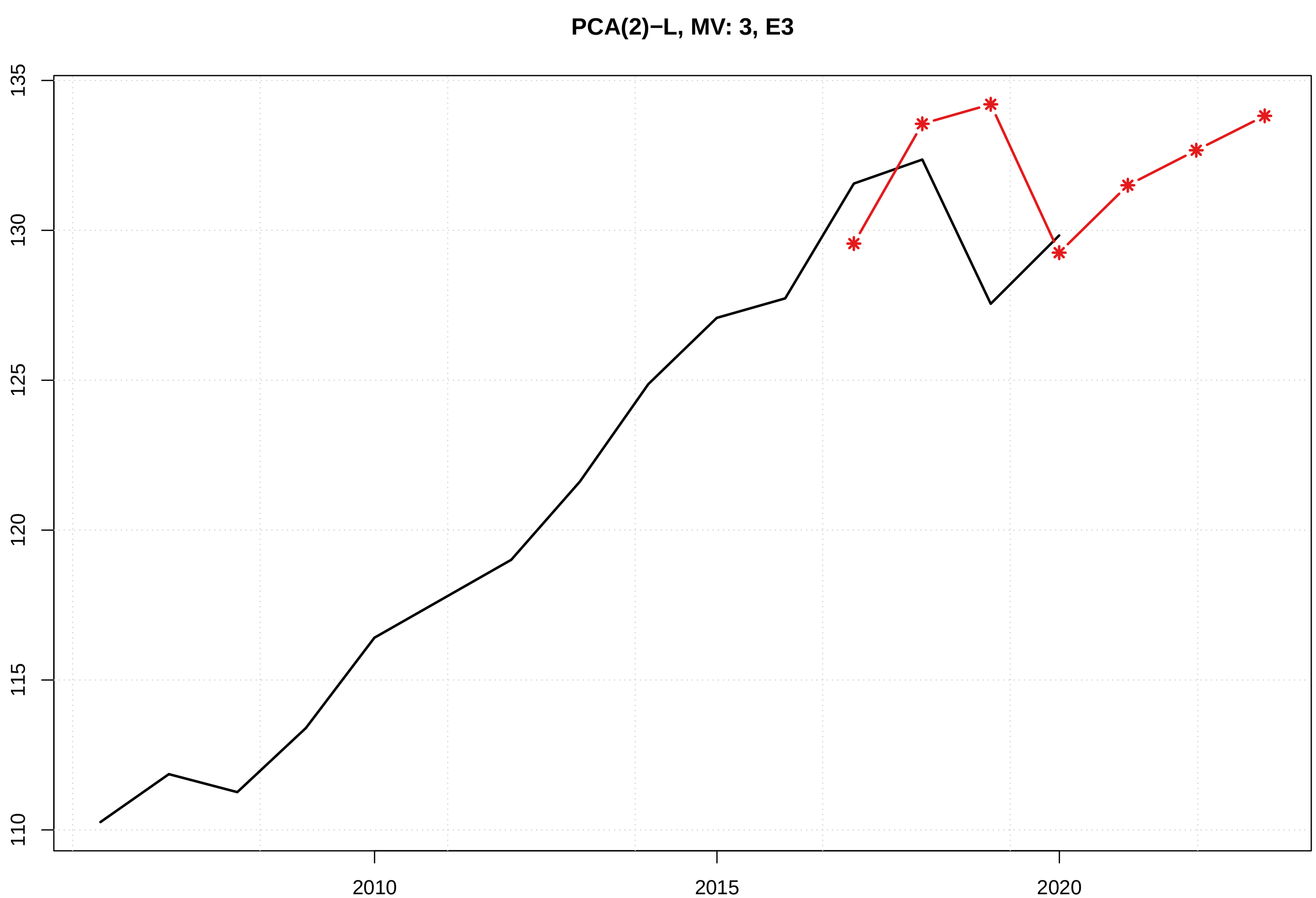
PCA(3)−S, MV: 3, I7



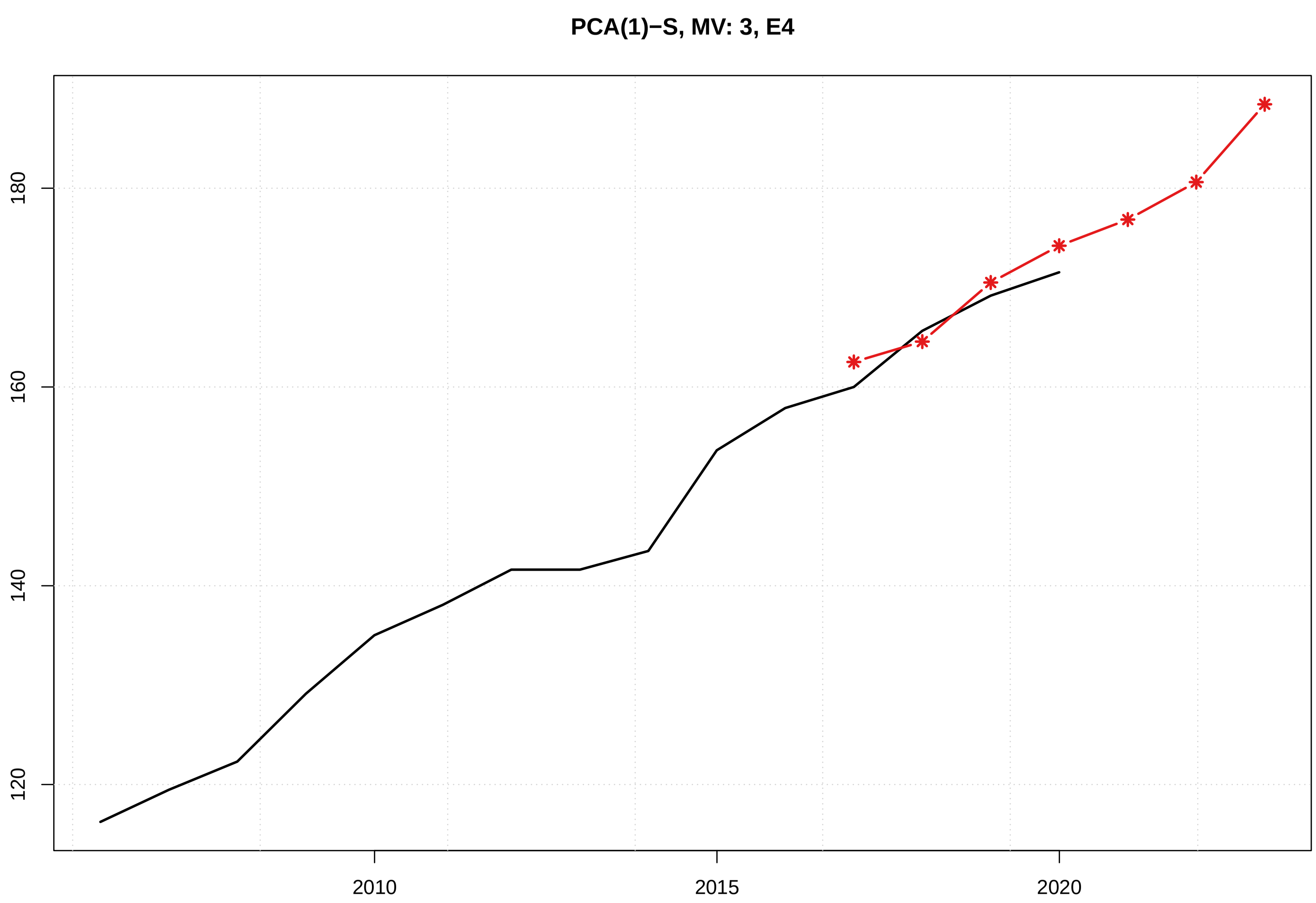

AdLasso−S, MV: 3, AGG
AGG
BottomUp
650000
600000
550000
500000
450000
400000
2010
2015
2020

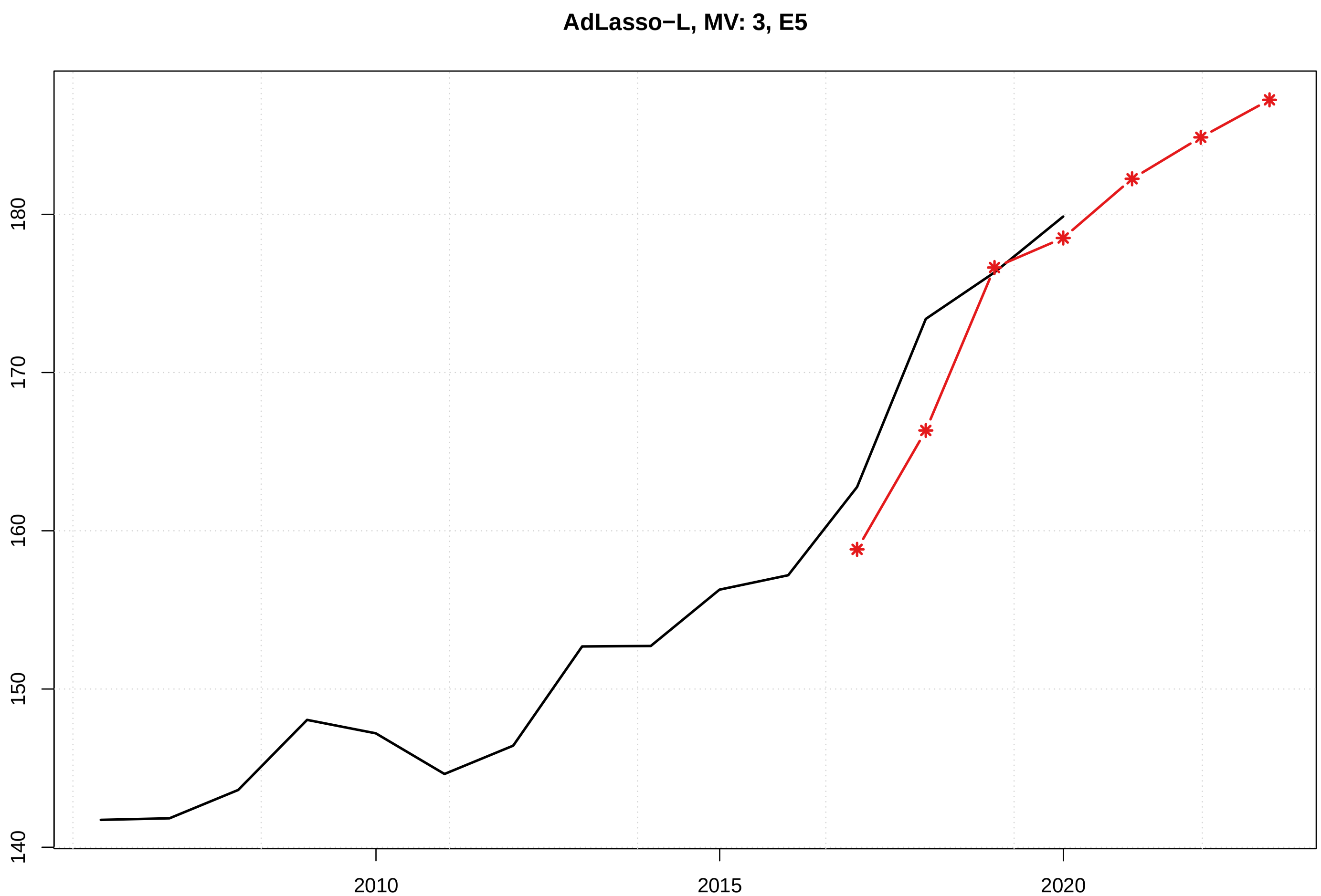

EN05−S, MV: 3, E1
170
160
150
140
130
2010
2015
2020


BFS−L, MV: 3, E2
160
150
140
130
2010
2015
2020

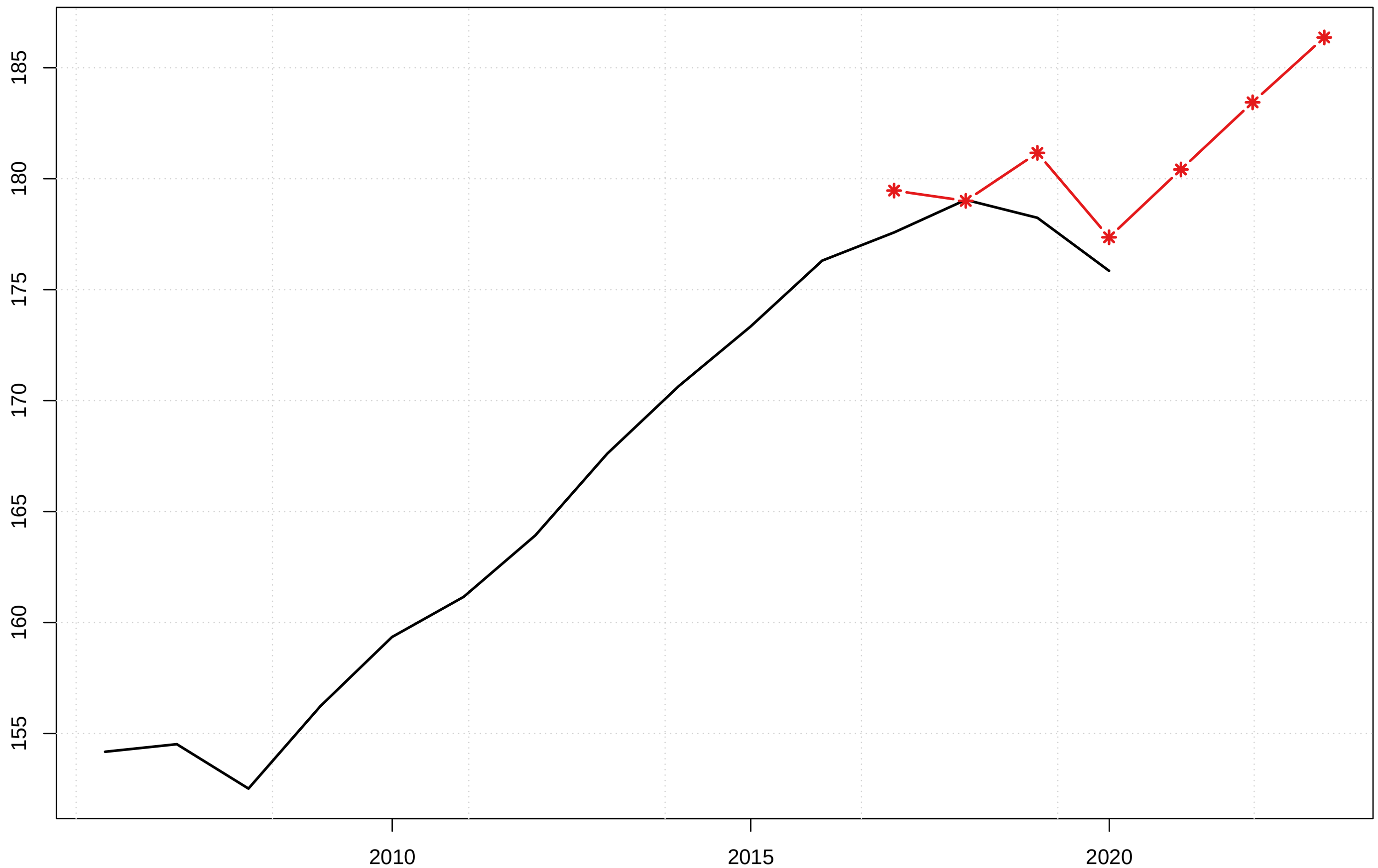


PCA(2)−L, MV: 3, E3
135
130
125
120
115
110
2010
2015
2020

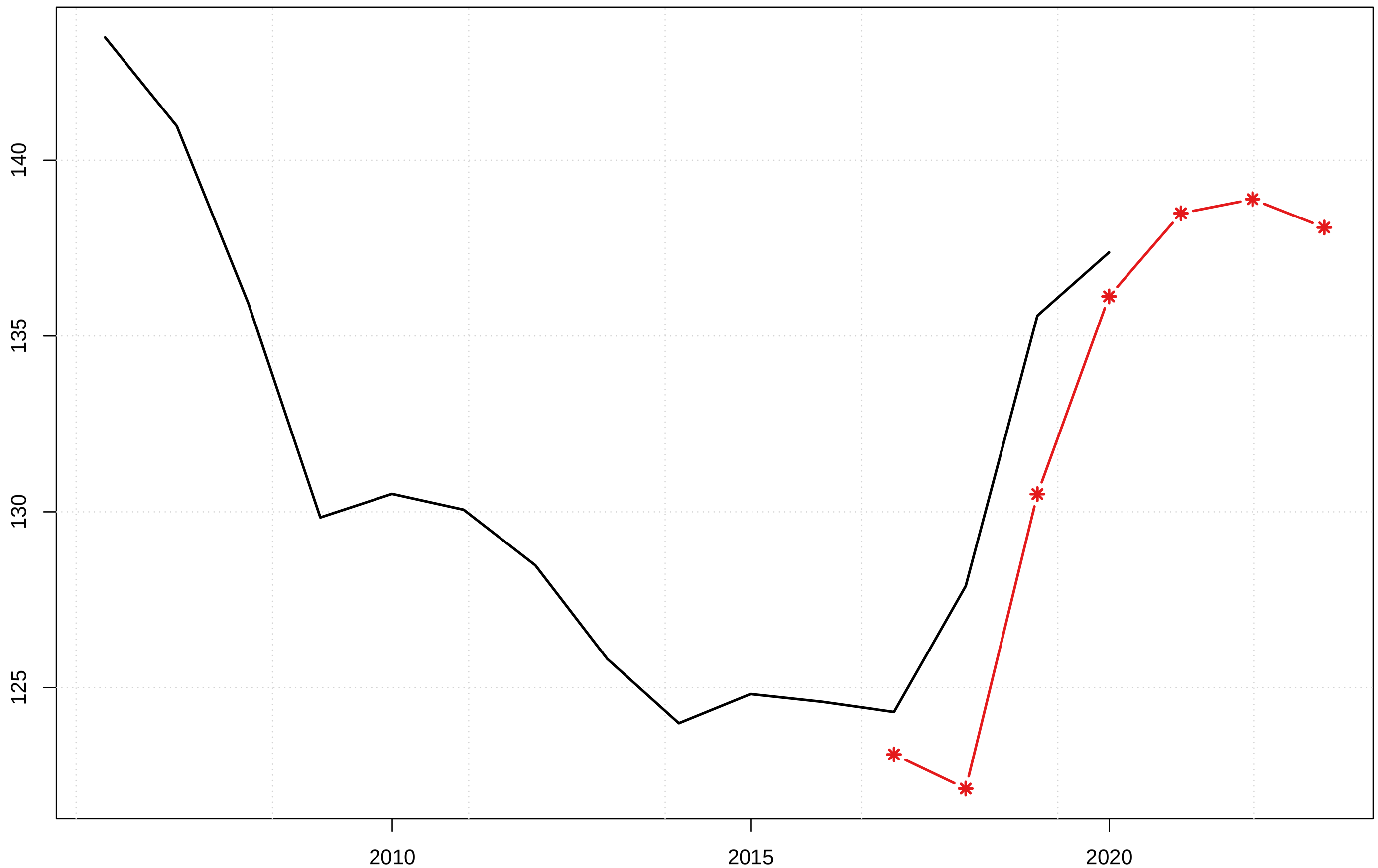

PCA(1)−S, MV: 3, E4
180
160
140
120
2010
2015
2020

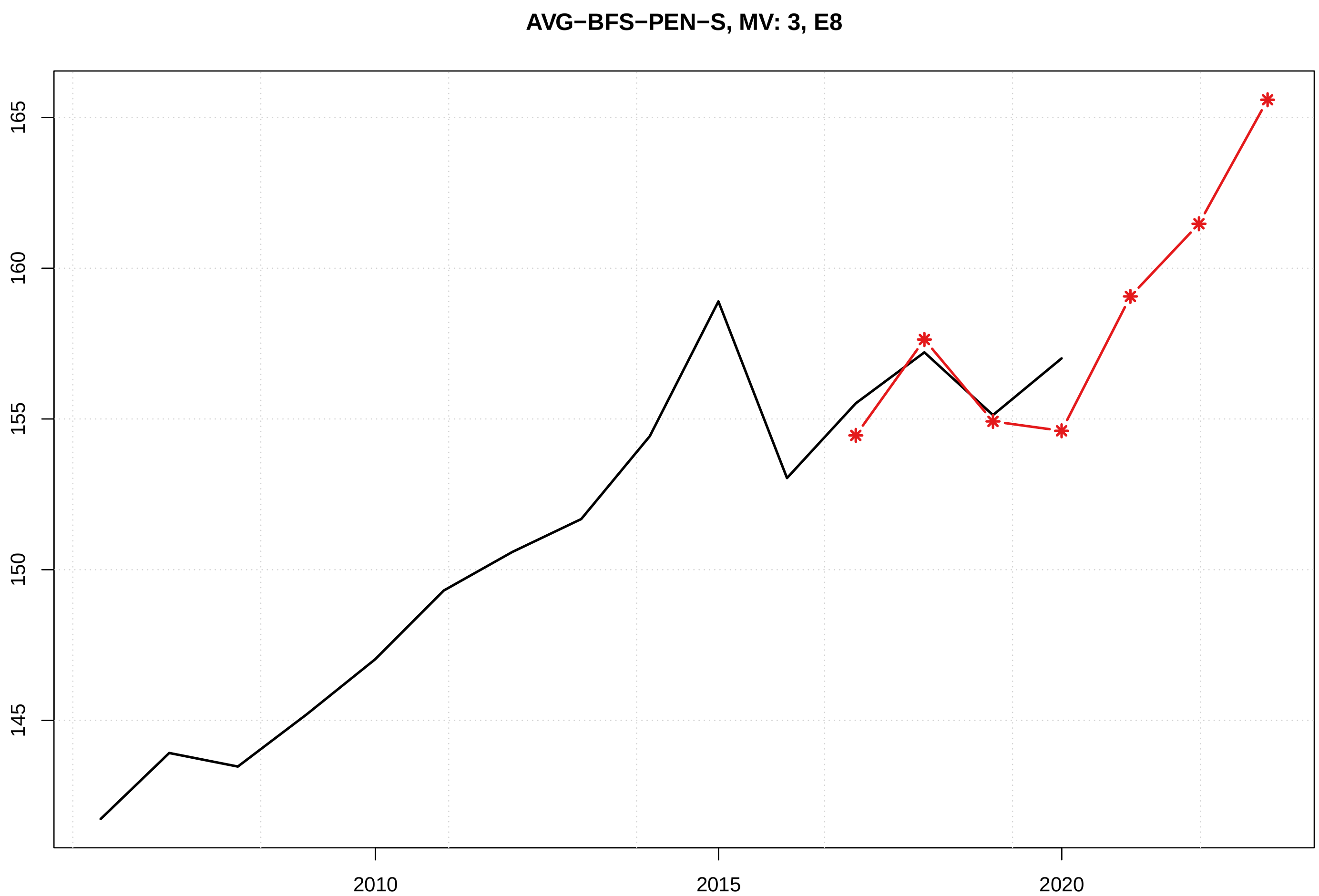

AdLasso−L, MV: 3, E5
140
150
160
170
180
2010
2015
2020

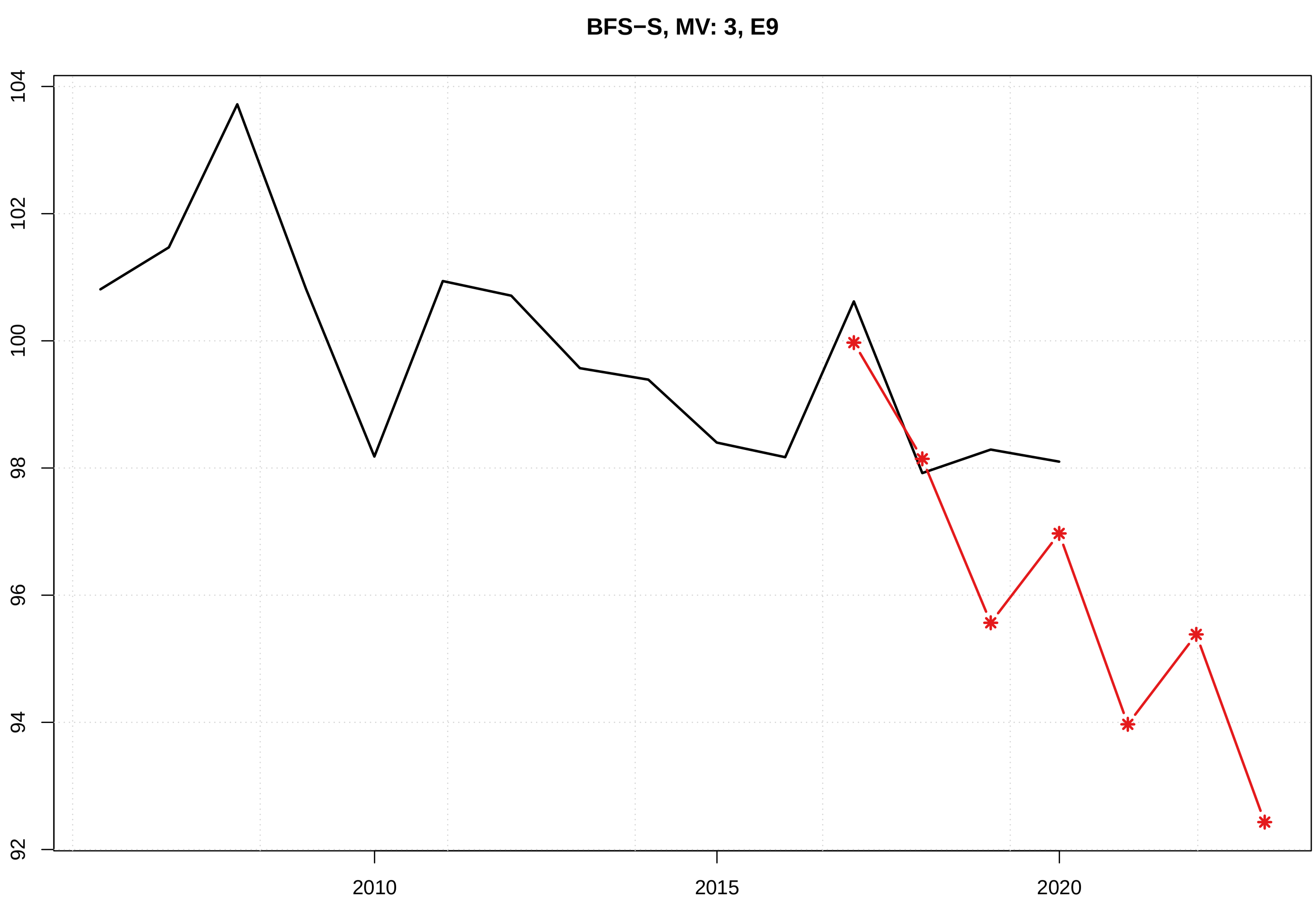
MLP−S, MV: 3, E6



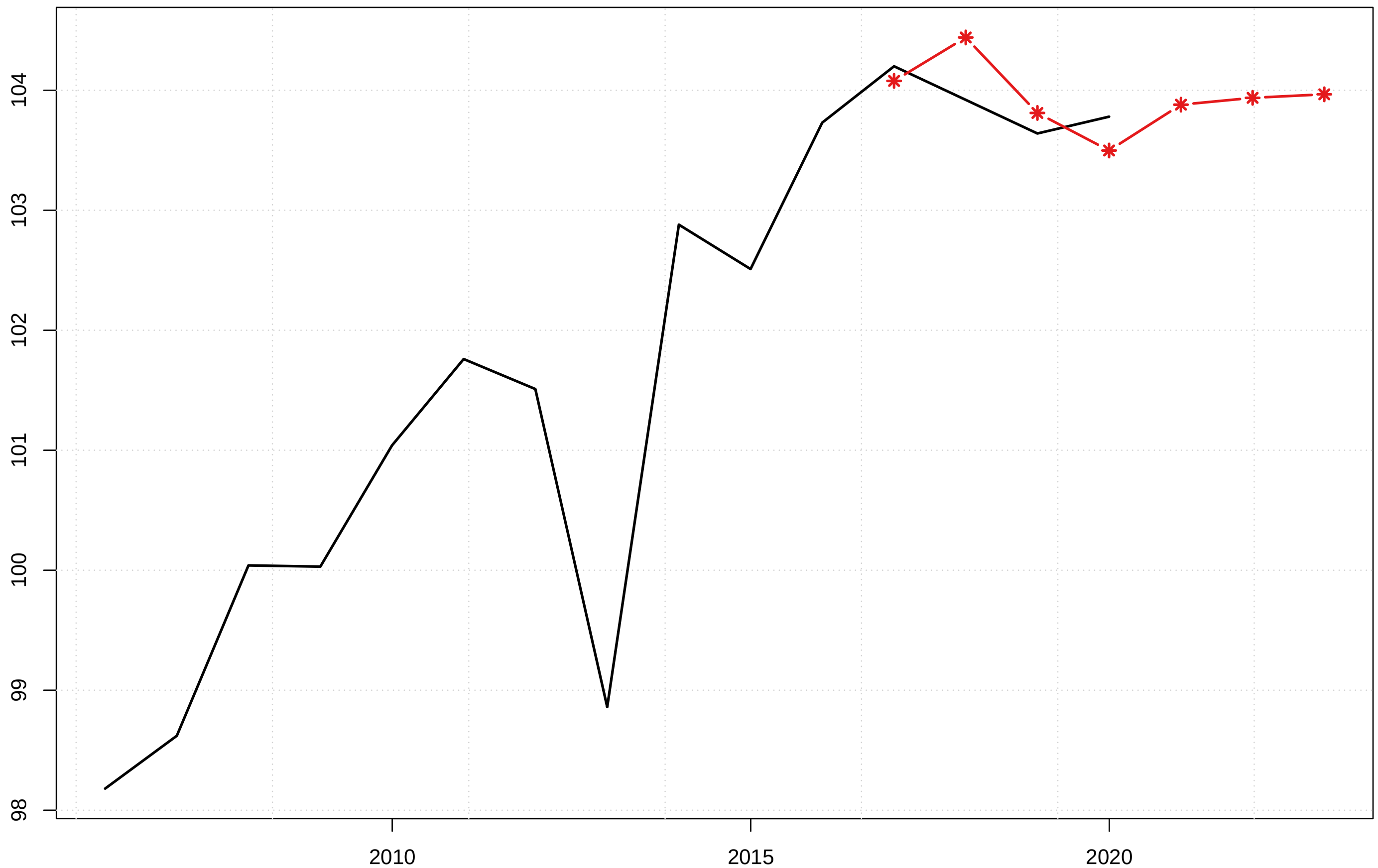
AdEN05−L, MV: 3, E7



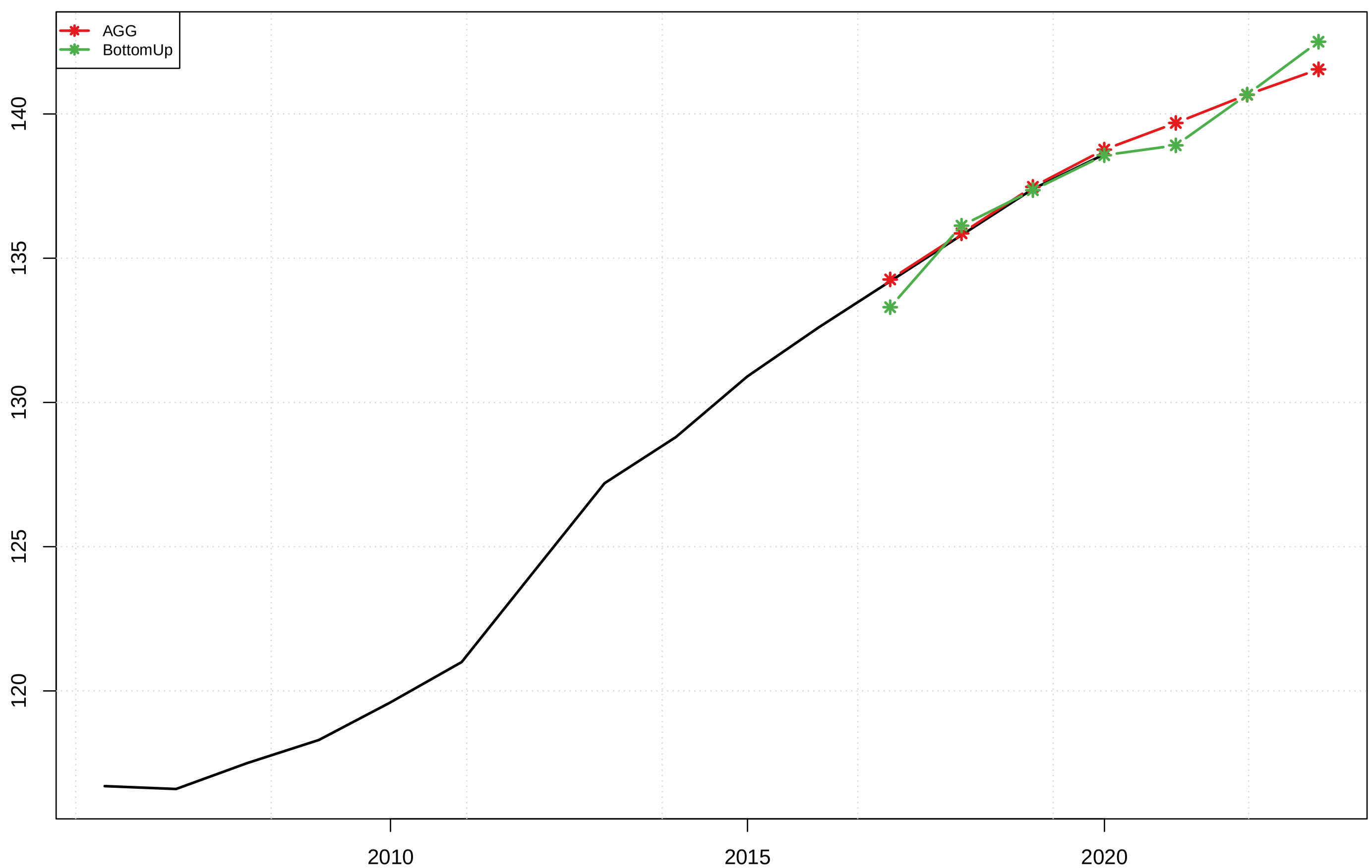

AVG−BFS−PEN−S, MV: 3, E8
165
160
155
150
145
2010
2015
2020

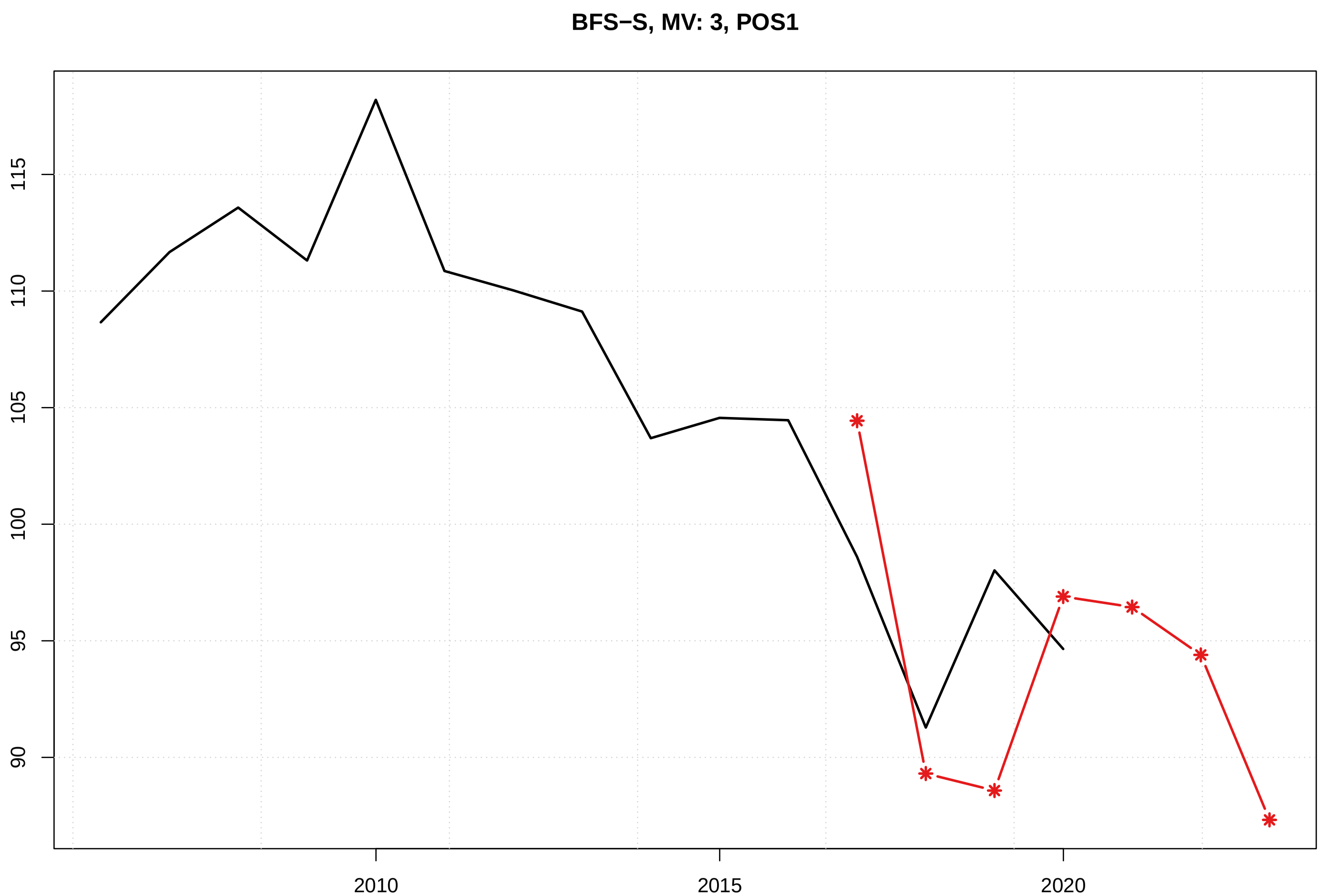

BFS−S, MV: 3, E9
104
102
100
98
96
94
92
2010
2015
2020

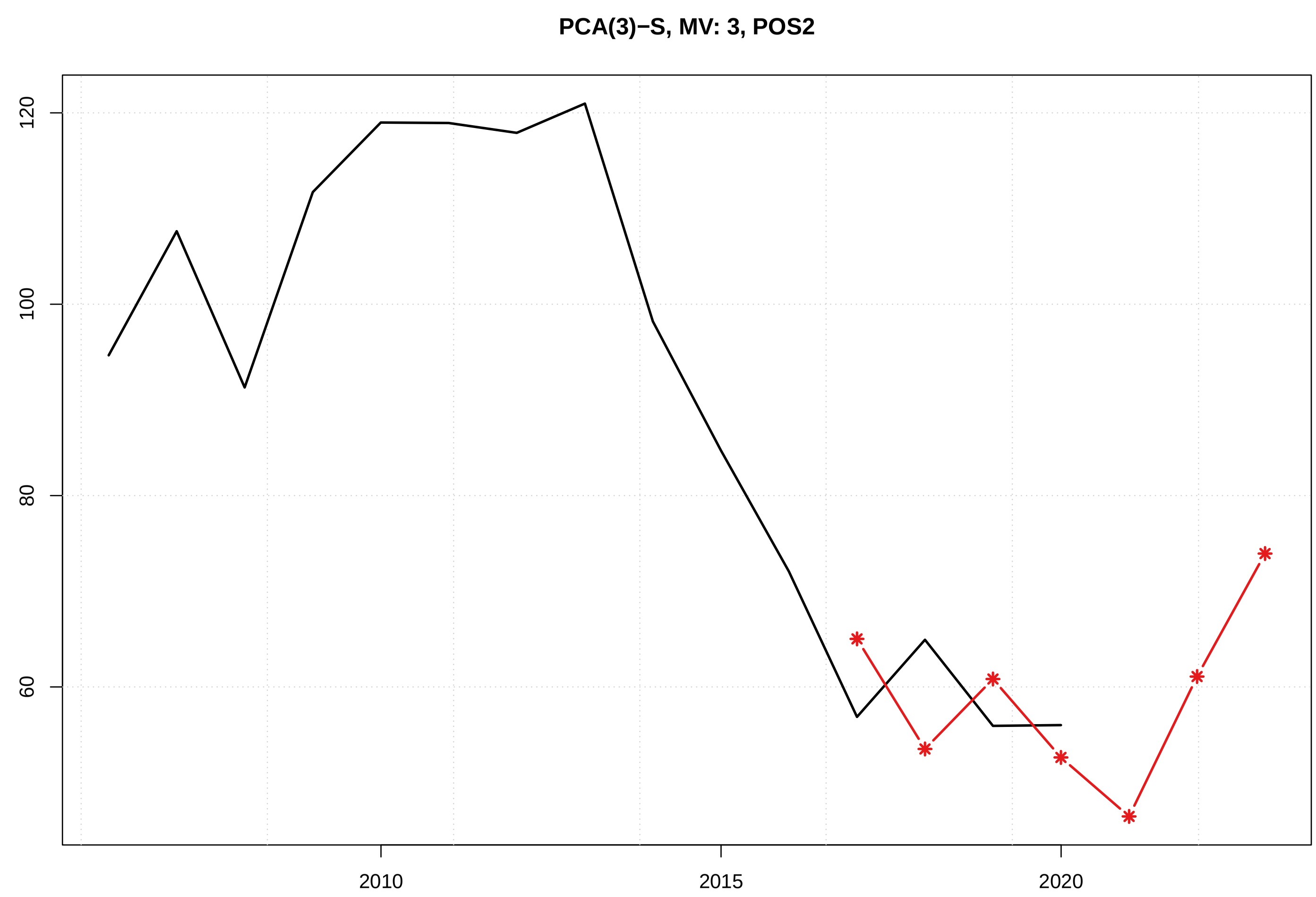
MLP−L, MV: 3, E10



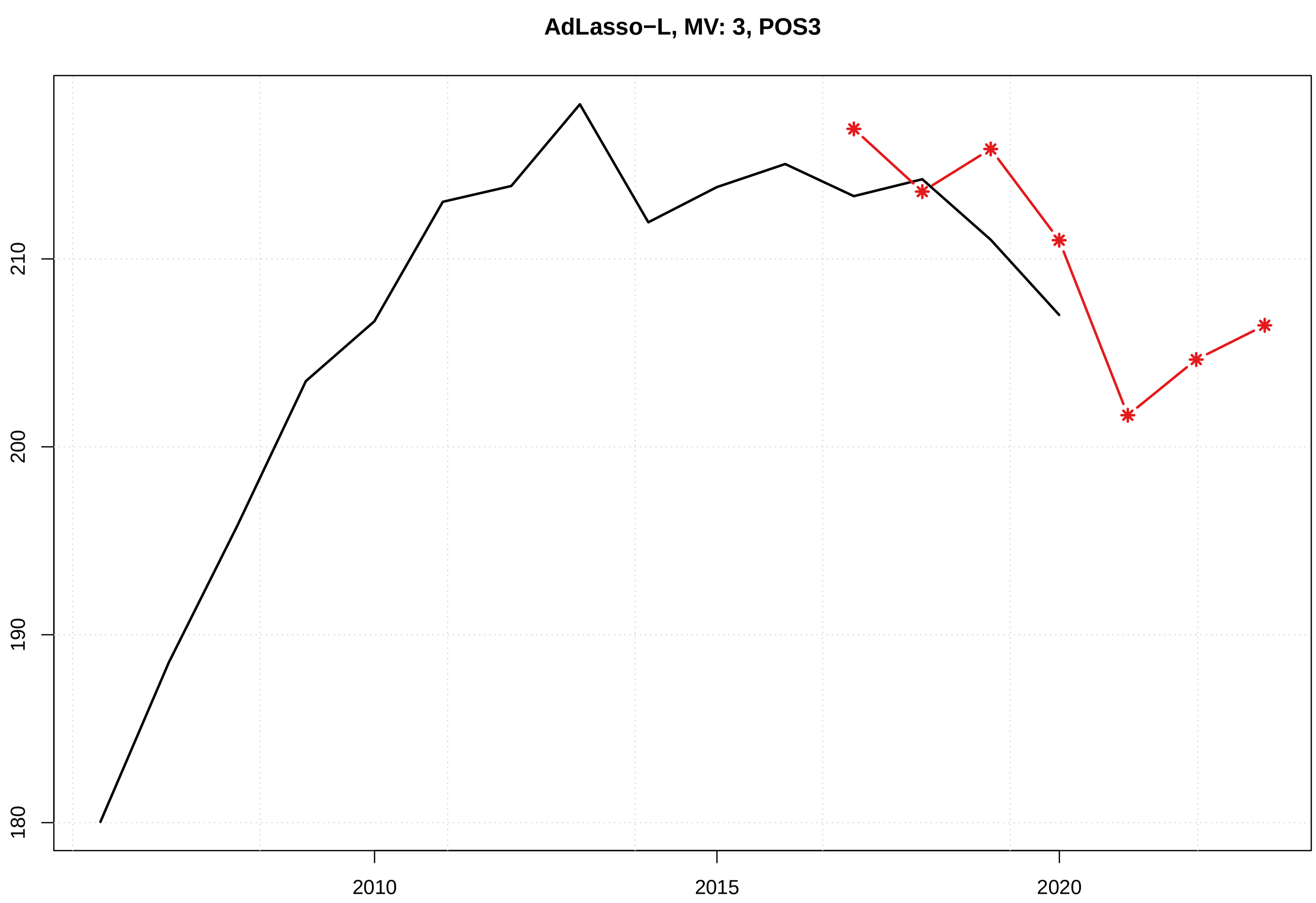
AVG−AR, MV: 3, AGG

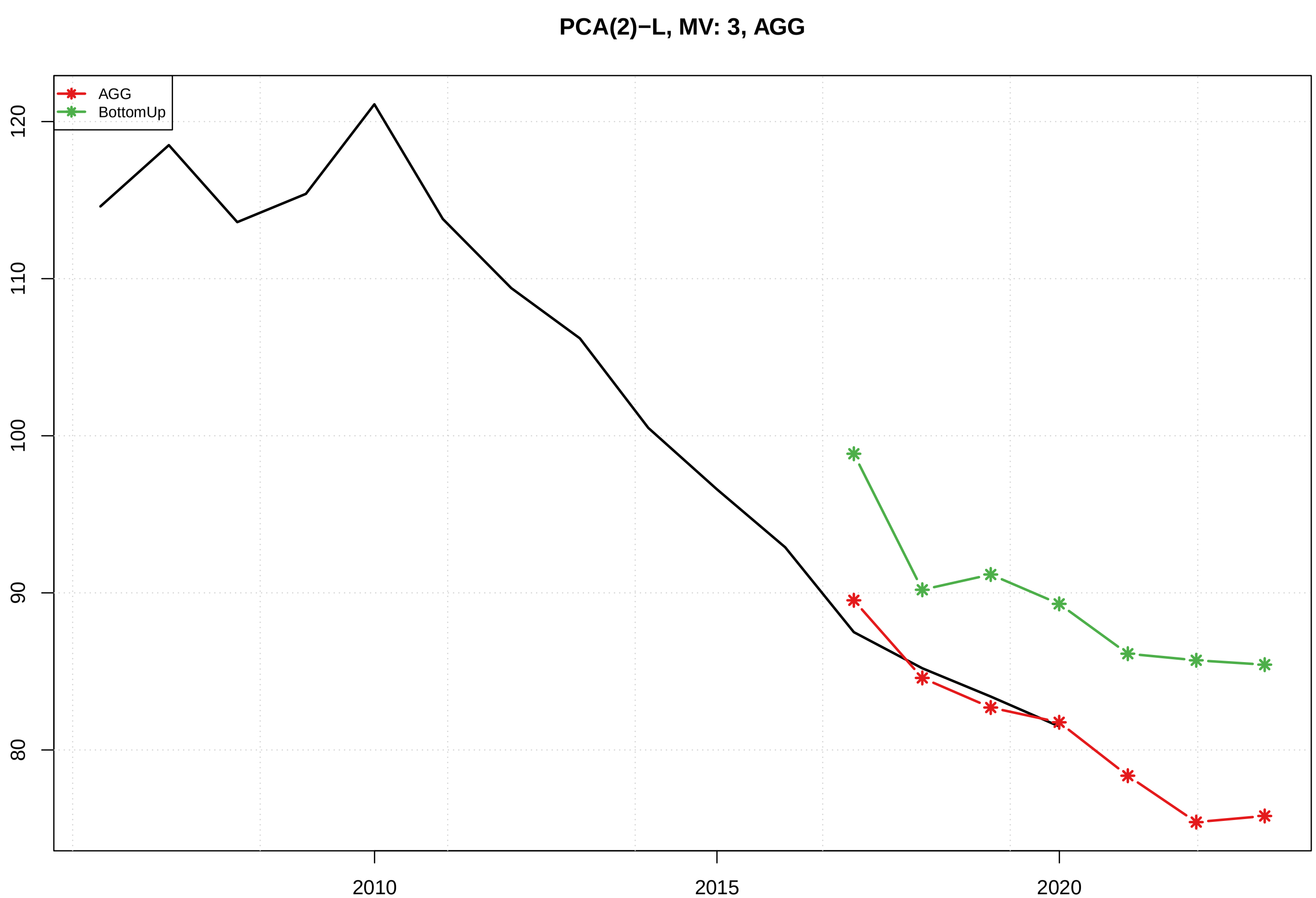



BFS−S, MV: 3, POS1
115
110
105
100
95
90
2010
2015
2020



PCA(3)−S, MV: 3, POS2
120
100
80
60
2010
2015
2020


AdLasso−L, MV: 3, POS3
210
200
190
180
2010
2015
2020



PCA(2)−L, MV: 3, AGG
AGG
BottomUp
120
110
100
90
80
2010
2015
2020