



**HAL**  
open science

## Backcasting world trade growth using data reduction methods

Amélie Charles, Olivier Darné

► **To cite this version:**

Amélie Charles, Olivier Darné. Backcasting world trade growth using data reduction methods. *The World Economy*, 2022, 45 (10), pp.3169 - 3191. 10.1111/twec.13274 . hal-04027843

**HAL Id: hal-04027843**

**<https://hal.science/hal-04027843>**

Submitted on 14 Mar 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Backcasting world trade growth using data reduction methods

Amélie Charles<sup>1</sup> | Olivier Darné<sup>2</sup> 

<sup>1</sup>Audencia Business School, Nantes  
Cedex 3, France

<sup>2</sup>LEMNA, University of Nantes, Nantes,  
France

## Correspondence

Olivier Darné, LEMNA, University of  
Nantes, Chemin de la Censive du Tertre,  
BP 52231, 44322 Nantes, France.  
Email: [olivier.darne@univ-nantes.fr](mailto:olivier.darne@univ-nantes.fr)

## Funding information

None.

## Abstract

This paper compares backcasting performance of models based on variable selection to dynamic factor model for backcasting the world trade growth rate with two months ahead. The variable selection models are specified by applying penalised regressions and an automatic general-to-specific procedure, using a large data set. A recursive forecast study is carried out to assess the backcasting performance by distinguishing crisis and non-crisis periods. The results show that, some selection-based models exhibit a good backcasting performance during both periods. The more accurate backcasts seem to be SCAD, adaptive Elastic-Net and adaptive SCAD during the global financial crisis (GFC) and COVID-19 crisis, whereas it seems rather Lasso, Elastic-Net, adaptive Lasso and DFM during the non-crisis period. Amongst the predictors for backcasting world trade growth, it appears that the index of global economic conditions proposed by Baumeister et al. (*The Review of Economics and Statistics*, 2020), the PMI indicator on new export orders in manufacturing sector and the MSCI world index are relevant.

## KEYWORDS

backcasting, general-to-specific approach, shrinkage methods, variable selection, world trade growth

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *The World Economy* published by John Wiley & Sons Ltd.

## 1 | INTRODUCTION

Tracking world trade is an important and widely analysed indicator by policymakers and analysts, owing to the strong association between trade and economic growth. Therefore, the accuracy of world trade forecasts is important to provide a reliable and early analysis of the ongoing trade developments. However, monitoring world trade in real time is challenging because of significant publication delays. For example, the OECD and the WTO publish quarterly index of world trade with one quarter lag (Guichard & Rusticelli, 2011; Martnez-Martn & Rusticelli, 2021).<sup>1</sup> The Netherlands Bureau for Economic Policy Analysis (CPB) also publishes a monthly index of world trade, which is built based on the trade series of 85 countries, covering around 97% of the world trade volume, but is available with a lag of 2 months.<sup>2</sup>

To the best of our knowledge, only few papers forecast world trade growth and have shown that factor models prove better accurate, where the factors summarise large data sets relevant in the determination of world trade growth by extracting few common factors (dimensionality reduction; Burgert & Dees, 2009; Guichard & Rusticelli, 2011; Jakaitiene & Dees, 2012; Martnez-Martn & Rusticelli, 2021).<sup>3</sup> However, factor models suffers from a weakness due to the fact that they are often seen as a black box, especially because the factors can be very difficult to interpret.

In this paper, to provide an economic interpretation, we propose a number of models designed to backcast monthly world trade growth, based on a large monthly data set, with a lead of 2 months with respect to the CPB index publication and compare their backcasting performance to dynamic factor models.<sup>4</sup> We focus on the CPB index because it is considered as the benchmark indicator for global trade and is available in a monthly frequency and earlier than the OECD and WTO indexes. The variable selection models are specified by using penalised regressions and an automatic general-to-specific procedure.<sup>5</sup> We focus on penalised regression methods, which is a generalisation of ordinary least squares estimation, with an additional term that penalises the size of regression coefficients. In doing so, it regularises the

---

<sup>1</sup>The OECD publishes a world trade index using data from national accounts whereas the WTO publishes an aggregate of several sub-indices based on export orders, international air freight, container shipping, automobile sales and production, electronic components and agricultural raw materials (WTO, 2016).

<sup>2</sup>The WTO also provides a composite indicator on trends in the trade volume of global merchandise but with a lag of three months. It relies on Hodrick–Prescott filtering to aggregate sub-indices (WTO, 2016).

<sup>3</sup>Burgert and Dees (2009) and Jakaitiene and Dees (2012) forecast CPB world trade using direct and bottom-up approaches—that is methods based on the aggregation of country-specific forecasts—and various short-term indicators relevant for the world economy (leading indicators, surveys, financial variables, manufacturing activity indicators, commodity prices and shipping indices). They show that factor models appear to perform the best. Guichard and Rusticelli (2011) compare the forecasting performance of bridge models and dynamic factor models to forecast OECD world trade growth using a large set of monthly indicators and show that the dynamic factor approach seems to perform better. Recently, Martínez-Martín and Rusticelli (2021) analyse the performance of their world trade index, including hard and survey data along with leading indicators, in forecasting WTO world trade growth.

<sup>4</sup>Note that backcasting refers to predict past observations whereas nowcasting refers to current observations. We thank an anonymous referee for this remark.

<sup>5</sup>Considerable attention has focused on the forecasting of macroeconomic variables in a data-rich environment via the implementation of a variety of machine learning, variable selection and shrinkage methods (Kim and Swanson, 2014, 2018; Li and Chen, 2014; Smeekees and Wijler, 2018).

model complexity and avoids over-fitting that can cause the out-of-sample forecasting performance to deteriorate. Lasso (least absolute shrinkage and selection operator) regression method (Tibshirani, 1996) is likely the most well-known regularised regression method. However, Lasso regression has undesirable properties when  $N$  is greater than  $T$  or when there is a group of variables amongst which all pairwise correlations are very high (Zou & Hastie, 2005). Otherwise, Lasso can generate a selection bias between highly correlated variables, leading to arbitrary selection and under-representation of important variables. For that, we employ various penalised regressions to solve the variable selection problem of Lasso, namely adaptive Lasso (Zou, 2006), Elastic Net (EN, Zou & Hastie, 2005), SCAD (Fan & Li, 2001), adaptive EN (Zou & Zhang, 2009) and adaptive SCAD (Xiao & Xu, 2015). One of the advantages of these penalised regressions is to display a good trade-off between predictive accuracy and model interpretability.<sup>6</sup> For the econometric method, we use the automatic model selection procedure based on a general-to-specific (GETS) modelling strategy that allows the econometrician to exploit the availability of a large number of data (Doornik & Hendry, 2018). As shown by Castle (2005), GETS strategy is appropriate when there is a desire to conform to economic interpretation. Therefore, our paper appears to be the first to compare penalised regressions, GETS approach and dynamic factor model to backcast world trade growth rate. Further, we also extend the predictors employed in the previous studies by adding recent uncertainty indicators which can have an effect on world trade as well as the new index of global economic conditions proposed by Baumeister et al. (2020).

The remainder of the paper is organised as follows. Section 2 describes the data. Section 3 presents the methodology of the penalised regressions, automated GETS procedure and dynamic factor models, and the backcasting scheme is described in Section 4. Section 5 discusses the results of the variable selection and the out-of-sample backcasting exercise. Section 6 concludes.

## 2 | DATA

The data set includes monthly data for the monthly index of global trade published by the Netherlands Bureau for Economic Policy Analysis (CPB).<sup>7</sup> This index, which is currently considered the benchmark indicator for global trade, is available with a lag of two months.<sup>8</sup> The CPB index is built based on the trade series (prices and values) of 85 countries, covering around 97% of the world trade volume. Figure 1 shows that trade flows have known two large collapses during the Global Financial Crisis (GFC) and the COVID-19 crisis, with a sharp fall of 6.3% and 12.3% in November 2008 and April 2020 respectively (Figure 2).

<sup>6</sup>Interpretability is an advantage of penalized regressions over machine learning methods and factor models while predictive accuracy is in general more an advantage of penalized regression over OLS methods but not (systematically) over machine learning techniques and factor models.

<sup>7</sup>See van Welzenis and Suyker (2005) for a presentation of the CPB world trade series.

<sup>8</sup>The OECD publishes a index of world trade using data from national accounts (Guichard & Rusticelli, 2011; Martínez-Martín & Rusticelli, 2021) whereas the WTO publishes an aggregate of several sub-indices (WTO, 2016). However, these two indexes are at quarterly frequency and with a one quarter lag. Note that the CPB and WTO indexes refer to merchandise volumes, while the OECD index also includes services.

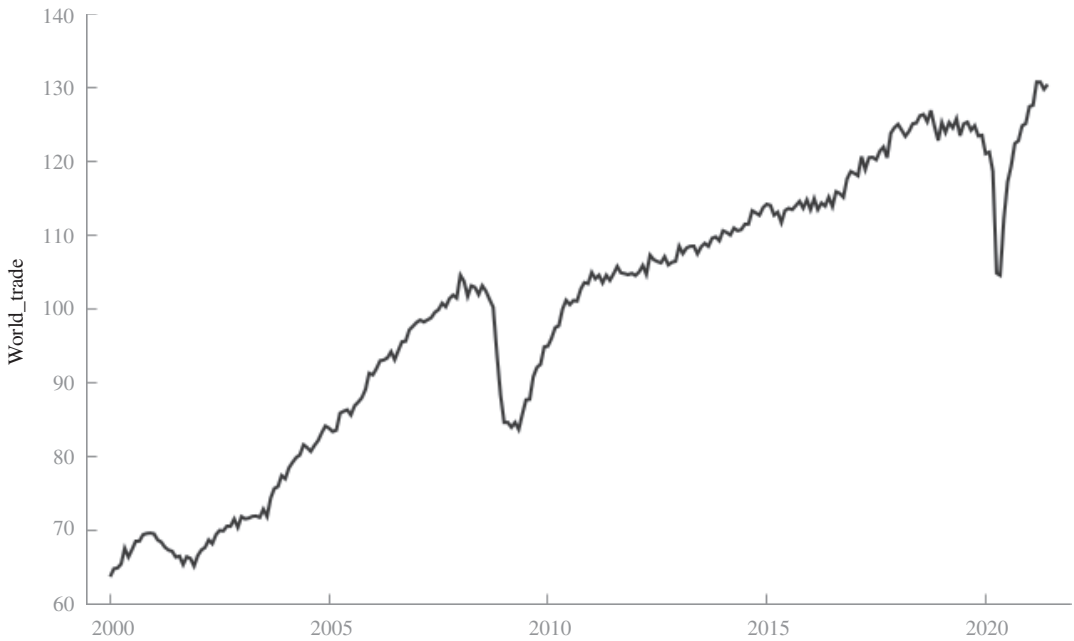


FIGURE 1 World trade in volume (monthly CPB index)—2000M1–2021M6

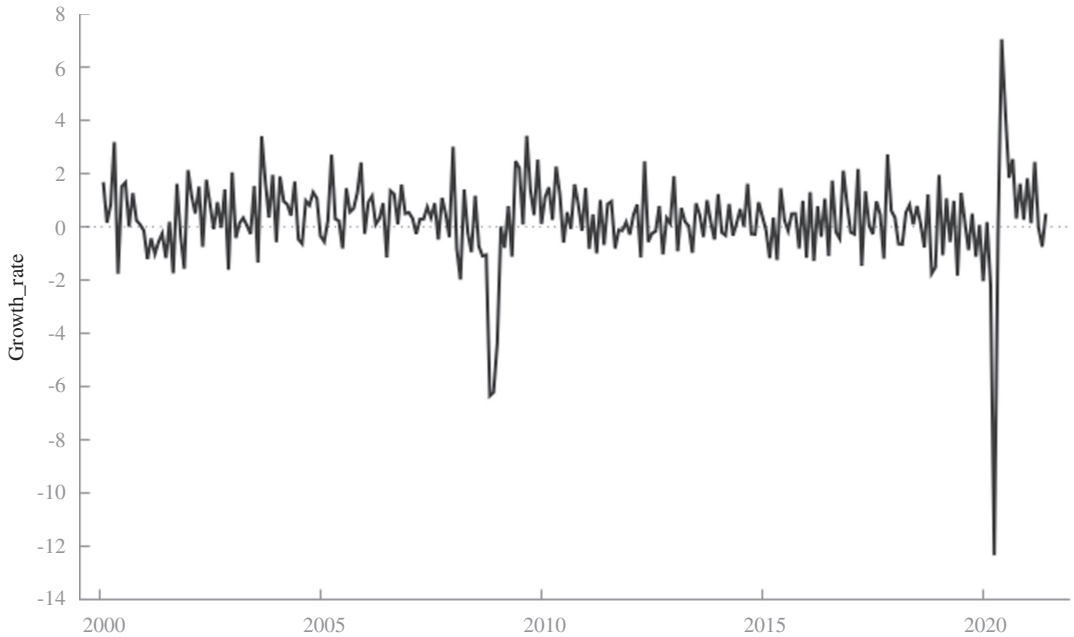


FIGURE 2 Growth rate of the monthly CPB index—2000M1–2021M6

We consider a large set of quantitative and qualitative indicators selected on the basis of their reliability and timeliness (i.e. 23 variables) covering the period of 2000M1–2021M6. These data are based on the predictors employed in Barhoumi et al. (2016), Guichard and Rusticelli (2011) and Martínez-Martín and Rusticelli (2021), namely surveys produced by the IHS Markit

and the IFO Institute (qualitative data), financial data on commodities, stock markets and exchange rates and maritime freight indexes. We also add uncertainty indicators which can have an effect on world trade with the Geopolitical risk (GPR) indicator of Caldara and Iacoviello (2018), the US trade policy uncertainty index which is one of the category-specific Economic Policy Uncertainty (EPU) indexes of Baker et al. (2016), and the trade policy uncertainty (TPU) index of Caldara et al. (2020). Finally, we add the new index of global economic conditions proposed by Baumeister et al. (2020) that covers a diverse range of variables tied to future energy demand. All the data are available with a publication lead of 2 months with respect to the CPB index in order to obtain a backcasting of the trade growth with a lead of 2 months. Note that the monthly financial data (e.g. Brent, CRB, BDI and GSCI,) are obtained by averaging the daily data, and the other monthly data (e.g. PMIs, IFOs, GECond and EPU) are available few days after the end of the month of interest. The data sources are presented in Appendix A.<sup>9</sup> We also consider twelve autoregressive terms and one lag for each predictor.

All variables are transformed to achieve stationarity and their transformation is given in Table A1. For penalised regressions variables are also standardised to have mean 0 and unit standard deviation to ensure that the shrinkage is fair and to prevent the scale of the variables affecting the estimation.

### 3 | METHODOLOGY

#### 3.1 | Penalised regressions

We estimate several shrinkage estimators for linear models with the estimated coefficients given by

$$\hat{\beta} = \arg \min_{\beta_0, \dots, \beta_p} \left\{ \frac{1}{2} \sum_{i=1}^N \left( y_i - \sum_{j=1}^p \beta_j x_{i,j} \right)^2 + P(\lambda; \beta) \right\} \quad (1)$$

where  $P(\lambda; \beta)$  is a penalty function that depends on the penalty parameter (or tuning parameter)  $\lambda$ . We consider different choices for the penalty functions as described below.

##### 3.1.1 | Lasso regression

Lasso (*least absolute shrinkage and selection operator*) regression, which was introduced by Tibshirani (1996), is characterised by an  $L_1$  penalty function, allowing for sparsity:

$$P^{Lasso}(\lambda; \beta) = \lambda \sum_{j=1}^p |\beta_j| \quad (2)$$

<sup>9</sup>A number of predictors used in Guichard and Rusticelli (2011) and Martínez-Martín and Rusticelli (2021) are not taken into account because they are available in the same time of the CPB index, such as the industrial production indexes. Further, other uncertainty indicators and indicators of global economic activity are also not taken into account due their short span or their publication lags with respect to the CPB index.

Lasso regressions tend to produce estimated regression coefficients that are exactly zeros, and thus can be used for variable selection, where only predictors with nonzero estimates are considered to be important (sparse solution). However, Lasso regression has undesirable properties when  $N$  is greater than  $T$  or when there is a group of variables amongst which all pairwise correlations are very high (Zou & Hastie, 2005). Otherwise, Lasso can generate a selection bias between highly correlated variables.

### 3.1.2 | Elastic Net

Zou and Hastie (2005) develop the EN regularisation method, which is a generalisation including Lasso and Ridge as special cases, to avoid collinearity. The EN method simultaneously carries out automatic variable selection (as in Lasso by setting some coefficients to zero) and continuous shrinkage (as in Ridge), via a (convex) combination of both  $L_1$  and  $L_2$  penalty functions:

$$P^{EN}(\lambda; \beta) = \lambda \sum_{j=1}^p \left( \alpha \beta_j^2 + (1 - \alpha) |\beta_j| \right) \quad (3)$$

where  $\alpha \in [0; 1]$ . If  $\alpha = 0$  or  $\alpha = 1$  the EN regression is equivalent to Lasso or Ridge regression respectively. An advantage of the EN lies in its property of grouped selection, that is, the 'grouping effect'. In other words, a group of highly correlated variables tend to have coefficients of similar magnitude and be selected in the same time.

However, the Lasso and EN methods fail to achieve the oracle property because they estimate the larger nonzero coefficients with asymptotically non-ignorable bias (Zou & Hastie, 2005). The oracle property means that when the true parameters have some zero components, the regularisation parameters are estimated as 0 with probability tending to 1, and the nonzero components are estimated as well as when the correct submodel is known.<sup>10</sup>

### 3.1.3 | Adaptive Lasso regression

Zou (2006) proposes adaptive Lasso (aLasso) to solve the variable selection problem of Lasso, showing that the inclusion of some additional information regarding the importance of each variable could considerably improve the results. aLasso uses the same penalty as Lasso with the inclusion of a weighting parameter that comes from a first-step model that can be Lasso, Ridge or even OLS:

$$P^{aLasso}(\lambda; \beta) = \lambda \sum_{j=1}^p \hat{\omega}_j |\beta_j| \quad (4)$$

where  $\hat{\omega} = |\hat{\beta}_j^*|^{-\gamma}$ , with  $\gamma > 0$  and  $\hat{\beta}_j^*$  are the coefficients from the first-step model.

<sup>10</sup>Fan and Peng (2004) highlighted the importance of studying the validity of the oracle property.

### 3.1.4 | Adaptive Elastic Net

Zou and Zhang (2009) extend aLasso to adaptive EN (aEN) that combines the strengths of the quadratic regularisation and the adaptively weighted Lasso shrinkage. Under weak regularity conditions, they establish the oracle property of the aEN.<sup>11</sup>

$$P^{aEN}(\lambda; \beta) = \lambda \sum_{j=1}^p \left( \alpha \beta_j^2 + (1 - \alpha) \hat{\omega}_j |\beta_j| \right) \quad (5)$$

where  $\hat{\omega}_j = |\hat{\beta}_j|^{-\lambda}$  are the adaptive weights, with  $\lambda > 0$ , and  $\hat{\beta}_j = \hat{\beta}_j^{EN}$  are the EN estimators from the first-step model. If  $\alpha = 0$  then the aEN reduces to the aLasso.

### 3.1.5 | SCAD regression

Fan and Li (2001) propose a non-concave penalty function referred to as the smoothly clipped absolute deviation (SCAD) which penalises small coefficients more severely than Lasso and penalises large coefficients less severely, implying more parsimonious models and bias reduction respectively. Therefore, the SCAD estimator has the desirable properties, including unbiasedness, sparsity and continuity as well as the oracle property (Fan & Peng, 2004).

The SCAD estimator is given by

$$\hat{\beta}_{scad} = \underset{\beta_0, \dots, \beta_p}{\operatorname{argmin}} \left\{ \frac{1}{2} \sum_{i=1}^N \left( y_i - \sum_{j=1}^p \beta_j x_{ij} \right)^2 + \sum_{j=1}^p P^{scad}(\lambda; |\beta_j|) \right\} \quad (6)$$

where  $P_\lambda^{scad}(\beta)$  is the penalty function defined on  $[0, \infty)$ , with  $a > 2$ , by

$$P^{SCAD}(\lambda; \beta) = \begin{cases} \lambda |\beta| & \text{if } |\beta| \leq \lambda \\ \frac{(2a\lambda|\beta| - \beta^2 - \lambda^2)}{2(a-1)} & \text{if } \lambda < |\beta| \leq a\lambda \\ \frac{\lambda^2(a^2 + 1)}{2} & \text{if } |\beta| > a\lambda \end{cases} \quad (7)$$

Xiao and Xu (2015) propose the adaptive version the SCAD regression, called aSCAD.<sup>12</sup>

<sup>11</sup>They show by simulations that the adaptive EN deals with the collinearity problem better than the other oracle-like methods (aLasso and SCAD), thus displaying much improved finite sample performance.

<sup>12</sup>Xiao and Xu (2015) develop the multi-step aEN and aSCAD, a multi-step estimation algorithm built upon aEN regularization for reducing the false positives in high-dimensional variable selection while still maintaining the estimation accuracy. We applied this method and the variable selection is similar to that of aEN and aSCAD for our data set.



### 3.2 | Automated GETS procedure

The automatic model selection procedure is based on a general-to-specific (GETS) modelling strategy, proposed by David Hendry.<sup>13</sup>

Autometrics is the computer implementation used in PcGive (Doornik & Hendry, 2018) and based on the automated GETS strategy. This automatic model selection procedure has three basic stages in its approach to select a parsimonious undominated representation of an overly general initial model, denoted the general unrestricted model (GUM) containing all variables likely (or specified) to be relevant, including the maximum lag length of the independent and dependent variables: (i) formulate the GUM that passes a set of chosen diagnostic tests.<sup>14</sup> Each non-significant regressor in the GUM constitutes the starting point of a backward elimination path, and a regressor is non-significant if the  $p$ -value of the two-sided  $t$ -test is lower than the chosen significance level  $\alpha$  (5% by default) (pre-search process); (ii) undertake backward elimination along multiple path by removing, one-by-one, non-significant regressors as determined by the chosen significance level  $\alpha$ . Each removal is checked for validity against the chosen set of diagnostic tests, and for parsimonious encompassing against the GUM; and (iii) select, among the terminal models, the specification with the best fit according to the BIC.

### 3.3 | Dynamic factor model

In the factor model framework, variables  $X_t$ , are represented as the sum of two mutually orthogonal unobservable components: the common component  $\chi_t$  and the idiosyncratic component  $\xi_t$ . For a given  $t$ ,  $t = 1, \dots, T$ , the static factor model is defined by:

$$X_t = \Lambda F_t + \xi_t, \quad (8)$$

where  $X_t = [x_{1t}, \dots, x_{nt}]'$  is a vector of  $n$  stationary time series and it is assumed that the series have zero mean and covariance matrix  $\Gamma(0)$ ,  $\Lambda$  is the loading matrix such that  $\Lambda = [\lambda_1, \dots, \lambda_n]'$ , the common components  $\chi_t = \Lambda F_t$  are driven by a small number  $r$  of factors  $F_t$  common to all the variables in the model such that  $F_t = [F_{1t}, \dots, F_{rt}]'$ , and  $\xi_t = [\xi_{1t}, \dots, \xi_{nt}]'$  is a vector of  $n$  idiosyncratic mutually uncorrelated components, driven by variable-specific shocks.

Different approaches have been proposed to take dynamics into account in modelling the factors  $F_t$  from dynamic factor models (DFM).<sup>15</sup> Here, we use the dynamic PCA in the frequency domain suggested by Forni et al. (2005), also called generalised dynamic factor (GDF) model.<sup>16</sup> They estimate the common factors based on generalised principal components in which

<sup>13</sup>An overview of the literature and the developments leading to GETS modelling in particular, is provided by Campos et al. (2005).

<sup>14</sup>The statistic tests are the Jarque-Bera normality test, and the Ljung-Box serial correlation test on the residual and the squared residuals.

<sup>15</sup>See Barhoumi et al. (2013) and Stock and Watson (2015) for a survey on DFMs.

<sup>16</sup>We also applied other approaches of DFMs with the static principal component analysis (PCA) with autoregressive model for the factors (Stock and Watson, 2002), the DFM based on a state-space representation (Doz et al., 2011) and the DFM estimated from quasi-maximum likelihood (Doz et al., 2012). The results have been lower than those obtained from the GDF model and are available from the authors upon request.

observations are weighted according to their signal-to-noise ratio. They proceed in two steps. First, the density spectral matrix of the common and idiosyncratic components  $\hat{\Sigma}_\chi(\theta)$  and  $\hat{\Sigma}_\xi(\theta)$  are estimated. Inverse Fourier transformation provides the time-domain autocovariances of the common and idiosyncratic components  $\hat{\Gamma}_\chi(k)$  and  $\hat{\Gamma}_\xi(k)$  for  $k$ lags. In a second step, they compute the  $r$ linear combinations of  $X_t$  that maximise the contemporaneous covariance explained by the common factors  $\hat{Z}'_j, \hat{\Gamma}_\chi(0)\hat{Z}_j$ , with  $j = 1, \dots, r$ . This optimisation problem can be reformulated as the generalised eigenvalue problem  $\hat{\Gamma}_\chi(0)\hat{Z}_j = \hat{\mu}_j\hat{\Gamma}_\xi(0)\hat{Z}_j$ , where  $\hat{\mu}_j$  denotes the  $j$ -th generalised eigenvalue and  $\hat{Z}_j$  its  $(n \times 1)$  corresponding eigenvectors. The factor estimates are obtained as  $F_t^{FHLR} = \hat{Z}'X_t$ , where  $\hat{Z} = (\hat{Z}_1, \dots, \hat{Z}_r)$  is the  $(n \times r)$ matrix of the eigenvectors.

## 4 | BACKCASTING SCHEME

Models have been constructed to estimate global trade growth rate, in anticipation of their official release. We carry out a recursive out-of-sample backcast. As shown in Figure 2, the GFC and the COVID-19 crisis have a strong impact on the world trade, and thus it is possible that some models are more adequate during crisis periods and other models during (stable) non-crisis periods.<sup>17</sup> Therefore, we evaluate out-of-sample performances during both periods.

The models are initially estimated on three periods, namely from January 2000 to July 2007 ( $T = 90$  observations), from January 2000 to November 2009 ( $T = 118$  observations) and from January 2000 to December 2019 ( $T = 239$  observations), with an expanding window approach, where the window size gradually increases of one month at each estimation. The out-of-sample backcasts are thus generated for three periods, August 2007 to November 2009 (GFC period), December 2009 to December 2019 (non-crisis period) and January 2020 to March 2021 (COVID-19 period), producing  $H = 28, 121$  and  $18$  backcasts respectively. Parameters are estimated at each step. We compare the forecasting accuracy of the models with an AR(1) model and the DFM based on the approach of Forni et al. (2005) since the previous studies have shown its good forecasting performance (Burgert & Dees, 2009; Guichard & Rusticelli, 2011; Jakaitiene & Dees, 2012).<sup>18</sup>

For each month  $t$ , we provide backcasts of the world trade growth rate,  $\hat{Y}_t$ , few days after the end of the month of interest when all the data are available with a lead of 2 months. To assess the predictive accuracy, we use the classical mean squared error (MSE) criterion defined by the following equation

$$MSE = \frac{1}{H} \sum_{t=1}^H (Y_t - \hat{Y}_t)^2 \quad (9)$$

where  $H = 28, 121, 18$  is the number of months considered in the recursive forecast exercise according to the crisis and non-crisis periods, and  $Y_t$  is the observed value as realised by the CPB.

<sup>17</sup>We thank an anonymous referee to suggest this analysis on non-crisis and crisis periods.

<sup>18</sup>We apply the automatic model selection procedure to select the number of factors  $r$ , by setting a maximum number for each specification ( $r = 5$ ). We did not use information criteria for the number of factors because these tests have been developed assuming that  $n$  and  $T$  tend towards infinity, an assumption not satisfied given the small size of our data set.

To evaluate the forecasting performance, we also use the metric of out-of-sample (OOS)  $R^2$  given by

$$R_{OOS}^2 = 1 - \frac{MSE_{model}}{MSE_{benchmark}} \quad (10)$$

where  $MSE_{model}$  and  $MSE_{benchmark}$  are the MSE of the given and benchmark models (AR model or DFM), respectively. The given model is more accurate than the benchmark if its  $R_{OOS}^2$  is positive.

Obviously, simply comparing MSE-values does not take into account the sample uncertainty underlying observed forecast differences. This is why we additionally applied the test based on superior predictive ability (SPA hereafter) proposed by Hansen (2005) which allows for multiple comparison. Let  $d_{0j,t} = L_{0,t} - L_{j,t}$ ,  $t = 1, \dots, n$  and  $j = 1, \dots, m$ , with  $j \in \mathcal{M}_0$ , and  $\mathcal{M}_0$  a collection of models, the relative performance of model  $j$  with respect to the benchmark model (indexed by 0), and  $L_{j,t}$  is a loss function. The null hypothesis is expressed with respect to the best alternative model,  $H_0: E[d_{0j,t}] \leq 0$  for all  $j \in \mathcal{M}_0$ . The corresponding test statistic is

$$T_{\max, \mathcal{M}} = \max_{j \in \mathcal{M}} \frac{n^{1/2} \bar{d}_j}{\hat{\omega}_j} \quad (11)$$

where  $\bar{d}_j = n^{-1} \sum_{t=1}^n d_{0j,t}$  and  $\hat{\omega}_j^2 = \widehat{var}(n^{1/2} \bar{d}_j)$  is a consistent estimator of the asymptotic variance.  $p$ -values for the test are obtained by bootstrap, that is the SPA consistent  $p$ -value, noted SPAC. A significant  $p$ -value (SPAC) means the null hypothesis that 'the benchmark model is not outperformed' is rejected.

Finally, we employ the model confidence set (MCS) procedure proposed by Hansen et al. (2011) to determine the set,  $\mathcal{M}^*$ , that consists of a subset of equivalent models in terms of predictive ability which are superior to the other competing models from a collection of models,  $\mathcal{M}_0$ . The MCS procedure yields a model confidence set,  $\widehat{\mathcal{M}}^*$ , that is a set of models constructed to contain the best models with a given level of confidence. This MCS allows to get several models displaying equivalent forecasting performance and therefore gives robustness to the forecasting exercise rather than to base the forecasting analysis only on one model. The  $t$ -statistic is defined as

$$T_{\max, \mathcal{M}} = \max_{i \in \mathcal{M}} t_i \text{ with } t_i = \frac{\bar{d}_i}{\sqrt{\widehat{var}(\bar{d}_i)}} \quad (12)$$

where  $\widehat{var}(\bar{d}_i)$  denotes the estimate of  $var(\bar{d}_i)$ ,  $\bar{d}_i = m^{-1} \sum_{j \in \mathcal{M}} \bar{d}_{ij}$ , and  $\bar{d}_{ij} = n^{-1} \sum_{t=1}^n d_{ij,t}$  with  $d_{ij,t} = L_{i,t} - L_{j,t}$  for all  $i, j \in \mathcal{M}_0$ , and  $L_{i,t}$  is a loss function (here MSE).<sup>19</sup> The  $t$ -statistic is associated with the null hypothesis  $H_{0, \mathcal{M}}: E(\bar{d}_i) = 0$  for all  $i \in \mathcal{M}$ , where  $\mathcal{M} \subset \mathcal{M}_0$ . If  $H_{0, \mathcal{M}}$  is accepted at level  $\alpha$  then the MCS is the set  $\widehat{\mathcal{M}}_{1-\alpha}^*$ .<sup>20</sup>

<sup>19</sup> $\bar{d}_{ij}$  measures the relative sample loss between the  $i$ -th and  $j$ -th models, while  $\bar{d}_i$  is the sample loss of the  $i$ -th model relative to the average across the models in  $\mathcal{M}$ .

<sup>20</sup>The MCS  $p$ -values are calculated using bootstrap implementation with 10,000 resamples (Hansen et al., 2011).

## 5 | EMPIRICAL RESULTS

### 5.1 | Variable selection

In the penalised regressions the tuning parameters are selected via cross-validation, which is a data-driven method that is designed to maximise the expected out-of-sample predictive accuracy. For the EN approach, the parameter  $\alpha$  is chosen sequentially with  $0 < \alpha < 1$ .<sup>21</sup> For the adaptive version of the Lasso, EN and SCAD methods the estimators  $\hat{\beta}_j$  in the first-step model—used to compute the adaptive weights  $\hat{\omega}_j$ —are the estimates from the Ridge regression, and the parameter  $\gamma$  is set to 0.5. For Autometrics we compared two model strategies: *Liberal* and *Conservative*, i.e., ‘target size’, which means ‘the proportion of irrelevant variables that survives the simplification process’ (Doornik, 2009) was set to 5% (*Liberal*) and 1% (*Conservative*).<sup>22</sup> The variable selections were similar for both strategies.

The results of variable selection are given in Table 1 for the different methods. Overall, the approaches select various predictors but all of them choose the index of global economic conditions (GECON), the PMI indicator on new export orders in manufacturing sector and the MSCI world index, suggesting that these variables can be important predictors for the world trade growth rate. The future energy demand appears thus to be an interesting predictor since the index of GECON covers a broad variety of predictor variables tied to future energy demand.<sup>23</sup> Another interesting finding is that uncertainty indicators (GPR, TPU and USTPU) are selected by most of the approaches. The results also show that some variables are selected to the time  $t$  (contemporaneous value) and other with one lag or both, suggesting that some indicators are informative on the evolution of the world trade growth with 2 months in advance (e.g. GECON, PMI on new export orders, MSCI ou GPR) and others with one month (e.g. Brent, TPU, USTPU or BDI). Finally, the penalised methods tend to select a high number of predictors, including information on surveys, commodities, stock markets, maritime freight and uncertainty. Note that Autometrics displays the more parsimonious model with four indicators.

Table 2 gives the loadings of the three factors of the DFM. The results show that the Factor 1 is associated with all the information contained in the surveys (PMI and IFO) and GECON, the Factor 2 with financial information (commodities, stocks and exchange rate) and the Factor 3 with freight and uncertainty information. When comparing with the variables deemed more important by penalised regressions only few are contained in the Factors 1 and 2. Note that a notably difference with the regularised regressions is that only the contemporaneous values enter in the DFM.

<sup>21</sup>We test several values of  $\alpha$ , from 0.1 to 0.9 with a step of 0.1 and choose the one that optimizes the cross-validation criterion.

<sup>22</sup>Epprecht et al. (2021) compared Lasso and aLasso to Autometrics from a simulation experiment and genomic data with these two strategies.

<sup>23</sup>This index of GECON is based on different categories: real economic activity, commodity prices, financial indicators, transportation, uncertainty, expectations, weather and energy-related measures. Baumeister et al. (2020) show that their indicator gives the most accuracy forecasts for forecasting the real price of Brent and global petroleum consumption.

TABLE 1 Results of variable selection—2000M1–2021M6

| Series                            | Lasso | EN | SCAD | aLasso | aEN | aSCAD | Autometrics |
|-----------------------------------|-------|----|------|--------|-----|-------|-------------|
| GECON                             | ×     | ×  | ×    | ×      | ×   | ×     | ×           |
| PMI manuf. new export orders      | ×     | ×  | ×    | ×      | ×   | ×     | ×           |
| PMI manuf. output                 | ×     | ×  |      | ×      | ×   | ×     |             |
| BDI                               |       | ×  | ×    | ×      | ×   |       |             |
| ClarkSea                          |       |    |      |        | ×   |       |             |
| Brent                             |       |    | ×    |        |     |       |             |
| SP GSCI                           |       |    | ×    |        | ×   | ×     |             |
| MSCI                              | ×     | ×  | ×    | ×      | ×   | ×     | ×           |
| Spread                            |       |    |      |        | ×   | ×     |             |
| GPR                               | ×     | ×  | ×    | ×      | ×   | ×     |             |
| Brent (1)                         | ×     | ×  | ×    | ×      | ×   | ×     |             |
| TPU (−1)                          |       |    | ×    |        | ×   | ×     |             |
| USTPU (−1)                        | ×     | ×  | ×    | ×      | ×   |       |             |
| PMI comp. new orders (−1)         |       |    | ×    |        | ×   |       |             |
| PMI manuf. input (−1)             |       |    |      |        | ×   |       |             |
| PMI manuf. new export orders (−1) |       |    |      |        | ×   | ×     |             |
| PMI manuf. stocks (−1)            | ×     | ×  | ×    | ×      | ×   | ×     |             |
| Spread (−1)                       |       |    | ×    |        |     |       |             |
| BDI (−1)                          | ×     | ×  | ×    | ×      | ×   | ×     |             |
| SP GSCI (−1)                      |       |    |      |        | ×   |       | ×           |
| MSCI (−1)                         | ×     | ×  |      | ×      | ×   |       |             |
| IFO climate (−1)                  | ×     | ×  | ×    | ×      | ×   |       |             |

Note: See 6 for information on the variables.

Abbreviations: BDI, Baltic Dry Index; ClarkSea, ClarkSea Index; GECON, global economic conditions indicator; GPR, geopolitical risk; MSCI, World MSCI Index; RBUS, USD nominal effective exchange rate; Spread, US treasury yield spread; TPU, trade policy uncertainty; USTPU, US trade policy uncertainty.

## 5.2 | Backcasting results

For the backcasting evaluation we compare the various approaches to the AR model and the DFM. Results in terms of MSE are presented in Tables 3 and 4 as well as the  $R^2_{OOS}$  computed with the benchmark models, the SPA  $p$ -values and the MCS  $p$ -values with their ranks for the non-crisis and crisis periods respectively.

### 5.2.1 | Non-crisis period

The results in terms of MSE show that Lasso, Elastic-Net and adaptive Lasso produce MSEs close to that of the DFM, ranging between 0.914 (Lasso) and 0.925 (aLasso and DFM), whereas the other models display higher MSE, with the highest MSE for SCAD (1.223), during the non-crisis period. This result is confirmed by the  $R^2_{OOS}(DFM)$  because these three penalised methods have

TABLE 2 Loadings of the DFM—2000M1–2021M6

| Series            | Factor 1 | Factor 2 | Factor 3 |
|-------------------|----------|----------|----------|
| Brent             | 0.138    | 0.406    | 0.037    |
| TPU               | 0.014    | 0.026    | −0.314   |
| GPR               | 0.012    | 0.037    | −0.126   |
| USTPU             | 0.010    | −0.060   | −0.364   |
| GECond            | 0.242    | 0.204    | 0.047    |
| PMI comp. input   | 0.223    | −0.171   | 0.026    |
| PMI comp. orders  | 0.301    | −0.046   | 0.000    |
| PMI comp. output  | 0.301    | −0.053   | 0.003    |
| PMI manuf. input  | 0.239    | −0.166   | 0.025    |
| PMI manuf. export | 0.307    | −0.085   | 0.006    |
| PMI manuf. orders | 0.304    | −0.018   | −0.04    |
| PMI manuf. output | 0.306    | −0.045   | −0.033   |
| PMI manuf. stocks | 0.234    | −0.241   | 0.092    |
| PMI manuf. index  | 0.312    | −0.092   | 0.000    |
| Spread            | −0.026   | 0.108    | −0.076   |
| SPGSCI            | 0.179    | 0.386    | −0.039   |
| BDI               | 0.046    | 0.114    | −0.577   |
| MSCIworld         | 0.108    | 0.372    | 0.109    |
| IFO climate       | 0.200    | −0.191   | 0.122    |
| IFO expectation   | 0.285    | −0.084   | 0.041    |
| ClarkSea          | 0.076    | −0.071   | −0.604   |
| RBUS              | −0.104   | −0.366   | −0.048   |
| CRB               | 0.126    | 0.395    | 0.028    |

Note: See 6 for information on the variables.

Abbreviations: BDI, Baltic Dry Index; ClarkSea, ClarkSea index; GECON, global economic conditions indicator; GPR, geopolitical risk; MSCI, World MSCI Index; RBUS, USD nominal effective exchange rate; Spread, US treasury yield spread; TPU, trade policy uncertainty; USTPU, US trade policy uncertainty.

a positive value close to zero. These models also exhibit better MSEs than that of the AR model with positive value of the  $R^2_{OOS}(AR)$ . Further, the SPA test indicates that these three regularised models display an equal predictive ability to the DFM. Moreover, Lasso, Elastic-Net, adaptive Lasso and DFM exhibit MCS  $p$ -values up to 0.85, giving thus similar backcast performance, where Lasso has the rank 1 and the lowest MSE.

We further provide time-series plots of the cumulative squared prediction error (CSPE) for each model to assess their backcasting performance over time. Figure 3 shows that the CSPE of the different methods displays similar and rather stable behavior.

## 5.2.2 | Crisis periods

SCAD, aEN and aSCAD display lowest MSEs during both crises, where the smallest MSEs are produced by aSCAD and SCAD during the GFC and the COVID-19 crisis respectively. This result

TABLE 3 Backcast performance during non-crisis period—2009M12–2019M12

| Models      | MSE (%) | $R^2_{OOS}(AR)$ | $R^2_{OOS}(DFM)$ | SPA    | MCS   | Rank |
|-------------|---------|-----------------|------------------|--------|-------|------|
| Autometrics | 0.999   | -0.030          | -0.080           | 0.168  | 0.378 | 6    |
| Lasso       | 0.914   | 0.058           | 0.012            | 0.710  | 1.000 | 1    |
| EN          | 0.916   | 0.056           | 0.010            | 0.613  | 0.912 | 2    |
| SCAD        | 1.223   | -0.261          | -0.322           | 0.004* | 0.001 | 8    |
| aLasso      | 0.925   | 0.047           | 0.001            | 0.447  | 0.899 | 3    |
| aEN         | 1.096   | -0.130          | -0.184           | 0.003* | 0.001 | 8    |
| aSCAD       | 1.111   | -0.146          | -0.201           | 0.005* | 0.002 | 7    |
| DFM         | 0.925   | 0.046           | -                | 0.395  | 0.899 | 3    |
| AR(2)       | 0.970   | -               | -0.048           | 0.141  | 0.632 | 5    |

Notes:  $R^2_{OOS}(AR)$  and  $R^2_{OOS}(DFM)$  are the  $R^2_{OOS} = 1 - (MSE_{model}/MSE_{benchmark})$  where the benchmark model is the AR(2) model and the DFM model respectively. \* and \*\*Significant at the 5% and 10% level respectively, for the SPA<sub>c</sub> test, the null of the Hansen test for superior predictive ability. MCS denotes the  $p$ -value of the statistic  $T_{\max, \mathcal{M}}$  of Hansen et al. (2011) based on the MSE loss function. Rank gives model ranking position based on the  $\hat{\mathcal{M}}_{90\%}^s$ .

TABLE 4 Backcast performance during crisis periods—2007M8–2009M11 and 2020M1–2021M6

| Models         | MSE (%) | $R^2_{OOS}(AR)$ | $R^2_{OOS}(DFM)$ | SPA     | MCS   | Rank |
|----------------|---------|-----------------|------------------|---------|-------|------|
| 2007M8–2009M11 |         |                 |                  |         |       |      |
| Autometrics    | 3.170   | 0.435           | 0.045            | 0.229   | 0.583 | 4    |
| Lasso          | 4.212   | 0.249           | -0.268           | 0.106   | 0.114 | 7    |
| EN             | 3.797   | 0.323           | -0.143           | 0.109   | 0.186 | 6    |
| SCAD           | 3.107   | 0.446           | 0.064            | 0.130   | 0.738 | 3    |
| aLasso         | 4.997   | 0.109           | -0.505           | 0.101   | 0.097 | 8    |
| aEN            | 2.183   | 0.611           | 0.343            | 0.591   | 0.978 | 2    |
| aSCAD          | 2.158   | 0.615           | 0.350            | 0.605   | 1.000 | 1    |
| DFM            | 3.321   | 0.408           | -                | 0.166   | 0.214 | 5    |
| AR(2)          | 5.608   | -               | -0.689           | 0.086** | 0.062 | 9    |
| 2020M1–2021M6  |         |                 |                  |         |       |      |
| Autometrics    | 6.994   | 0.497           | 0.279            | 0.107   | 0.338 | 4    |
| Lasso          | 8.706   | 0.374           | 0.103            | 0.198   | 0.247 | 5    |
| EN             | 8.947   | 0.357           | 0.078            | 0.254   | 0.238 | 6    |
| SCAD           | 4.713   | 0.661           | 0.514            | 0.913   | 1.000 | 1    |
| aLasso         | 11.71   | 0.158           | -0.207           | 0.156   | 0.067 | 8    |
| aEN            | 5.604   | 0.597           | 0.422            | 0.137   | 0.605 | 2    |
| aSCAD          | 6.196   | 0.555           | 0.361            | 0.181   | 0.514 | 3    |
| DFM            | 9.703   | 0.303           | -                | 0.193   | 0.136 | 7    |
| AR(2)          | 13.91   | -               | -0.434           | 0.118   | 0.004 | 9    |

Notes:  $R^2_{OOS}(AR)$  and  $R^2_{OOS}(DFM)$  are the  $R^2_{OOS} = 1 - (MSE_{model}/MSE_{benchmark})$  where the benchmark model is the AR(2) model and the DFM model respectively. \* and \*\*Significant at the 5% and 10% level respectively, for the SPA<sub>c</sub> test, the null of the Hansen test for superior predictive ability. MCS denotes the  $p$ -value of the statistic  $T_{\max, \mathcal{M}}$  of Hansen et al. (2011) based on the MSE loss function. Rank gives model ranking position based on the  $\hat{\mathcal{M}}_{90\%}^s$ .

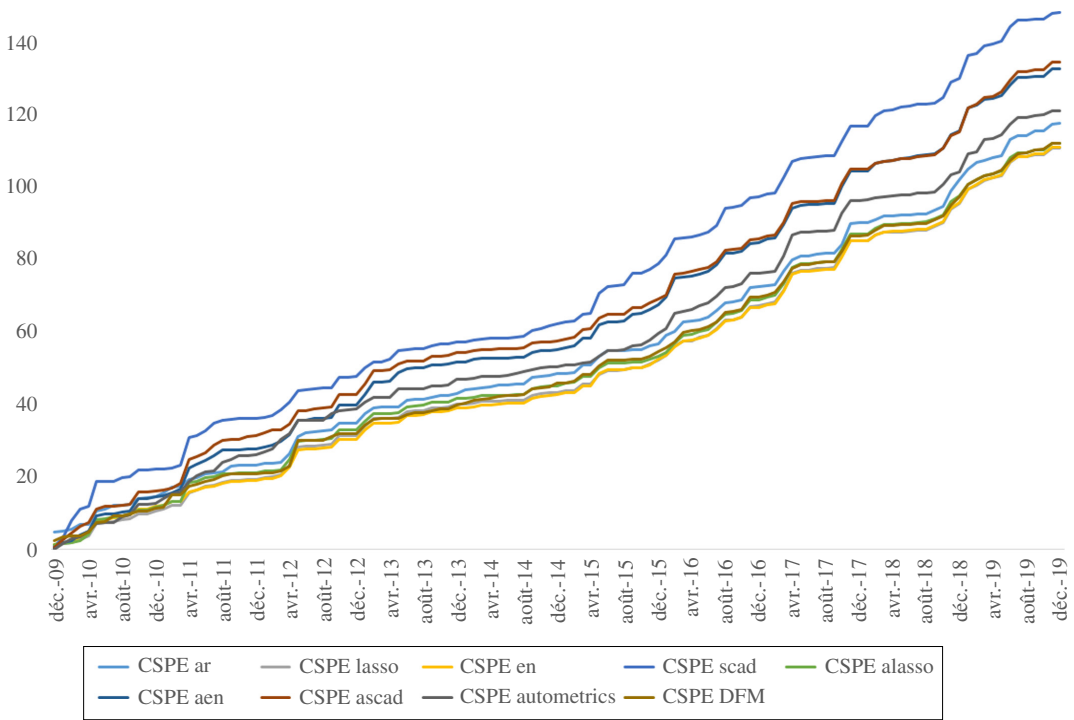


FIGURE 3 Cumulative squared prediction errors for the 2009M12–2019M12 period [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

is confirmed by their large positive values of  $R_{OOS}^2$  against both benchmarks and their high MCS  $p$ -values up to 0.70 and 0.50 for the GFC and the COVID-19 crisis, respectively, with the rank 1–3. These three models display a higher number of variables than the other models, suggesting that models having a better backcasting performance in crisis times are those accommodating for more variables (Table 5). Autometrics exhibits relevant backcasting with positive  $R_{OOS}^2$  and MCS  $p$ -values up to 0.30 for both crisis periods. Note that the AR(2) is significantly outperformed during the GFC crisis time.

Figures 4 and 5 show that the CSPE are very close between the various approaches until a strong break in December 2008 and in April 2020 during the GFC and the COVID-19 crisis, respectively. This break affects differently the models and can explain the difference in their backcasting performance. Table 5 displays the variable selection in December 2008 and in April 2020 and shows that the good backcasting performance of SCAD, aEN and aSCAD can be due to a higher number of variables and to some variables only selected by these methods, such as Brent, SP GSCI, Spread, TPU with one lag and the PMI indicator on new export orders in manufacturing sector with one lag.<sup>24</sup> Note that after few months the models exhibit a similar CSPE behavior.

Overall, these findings are interesting because some models based on variable selection exhibit both a good backcasting performance and an economic interpretation to the backcasts for the world trade growth, especially during crisis periods. Another interesting result is that

<sup>24</sup>Note that most of the methods select some identical predictors in December 2008 and in April 2020, such as GECON and PMI on new export orders in contemporaneous values and GECON, USTPU, PMI on stocks, BDI and SP GSCI with one lag, whereas some variables are only selected in December 2008, such as ClarkSea in contemporaneous and lag values, GPR and PMI on new export orders with one lag.



TABLE 5 Results of variable selection in December 2008 and April 2020

| Series                      | December 2008 |    |      |        |     |       | April 2020  |       |    |      |        |     |       |             |
|-----------------------------|---------------|----|------|--------|-----|-------|-------------|-------|----|------|--------|-----|-------|-------------|
|                             | Lasso         | EN | SCAD | aLasso | aEN | aSCAD | Autometrics | Lasso | EN | SCAD | aLasso | aEN | aSCAD | Autometrics |
| World trade (-2)            |               |    | X    |        | X   |       |             |       |    |      |        |     |       |             |
| GECON                       | X             | X  | X    | X      | X   | X     |             | X     | X  | X    | X      | X   | X     | X           |
| PMIm new export orders      | X             | X  | X    | X      | X   | X     |             | X     | X  | X    | X      | X   | X     | X           |
| ClarkSea                    |               | X  | X    | X      | X   | X     | X           |       |    |      |        |     |       |             |
| Brent                       |               |    |      |        |     |       |             |       | X  |      |        |     |       |             |
| SP GSCI                     |               |    | X    |        |     |       |             |       | X  | X    | X      | X   | X     | X           |
| MSCI                        |               |    |      |        |     |       |             |       | X  | X    | X      | X   | X     | X           |
| Spread                      |               |    | X    |        |     |       |             |       | X  | X    | X      | X   | X     | X           |
| TPU                         |               | X  | X    |        | X   |       | X           |       |    |      | X      |     |       |             |
| USTPU                       |               |    | X    |        |     |       |             |       |    |      |        |     |       |             |
| GPR                         |               | X  | X    | X      | X   | X     | X           |       | X  | X    | X      | X   | X     | X           |
| GECON (-1)                  | X             | X  | X    | X      | X   | X     |             | X     | X  | X    | X      |     |       |             |
| Brent (1)                   |               |    |      |        |     |       |             | X     |    |      |        |     |       |             |
| TPU (-1)                    |               |    | X    |        |     |       |             |       |    | X    | X      | X   | X     | X           |
| USTPU (-1)                  |               | X  | X    |        | X   |       | X           |       | X  | X    | X      | X   | X     | X           |
| GPR (-1)                    |               | X  | X    | X      | X   | X     | X           |       | X  | X    | X      | X   | X     | X           |
| PMIm new orders (-1)        | X             | X  |      | X      | X   | X     |             |       |    |      |        |     |       |             |
| PMIm new export orders (-1) |               |    |      |        |     |       |             |       |    |      | X      | X   | X     | X           |

TABLE 5 (Continued)

| Series           | December 2008 |    |      |        |     |       | April 2020  |       |    |      |        |     |       |             |
|------------------|---------------|----|------|--------|-----|-------|-------------|-------|----|------|--------|-----|-------|-------------|
|                  | Lasso         | EN | SCAD | aLasso | aEN | aSCAD | Autometrics | Lasso | EN | SCAD | aLasso | aEN | aSCAD | Autometrics |
| PMIm output (-1) |               | X  |      |        | X   | X     | X           |       |    |      |        |     |       |             |
| PMIm stocks (-1) |               |    | X    |        | X   | X     | X           |       | X  |      | X      | X   | X     | X           |
| BDI (-1)         |               | X  |      | X      | X   |       |             |       | X  | X    |        |     |       |             |
| ClarkSea (-1)    |               | X  | X    | X      | X   | X     | X           |       |    |      |        |     |       |             |
| SP GSCI (-1)     |               | X  | X    | X      | X   | X     | X           | X     | X  | X    | X      | X   | X     | X           |
| MSCI (-1)        |               | X  | X    | X      | X   | X     | X           | X     | X  | X    | X      | X   | X     | X           |

Note: See 6 for information on the variables.

Abbreviations: BDI, Baltic Dry Index; ClarkSea, ClarkSea index; GECON, global economic conditions indicator; GPR, geopolitical risk; MSCI, World MSCI Index; PMIm, PMI in manufacturing sector; RBUS, USD nominal effective exchange rate; Spread, US treasury yield spread; TPU, trade policy uncertainty; USTPU, US trade policy uncertainty.

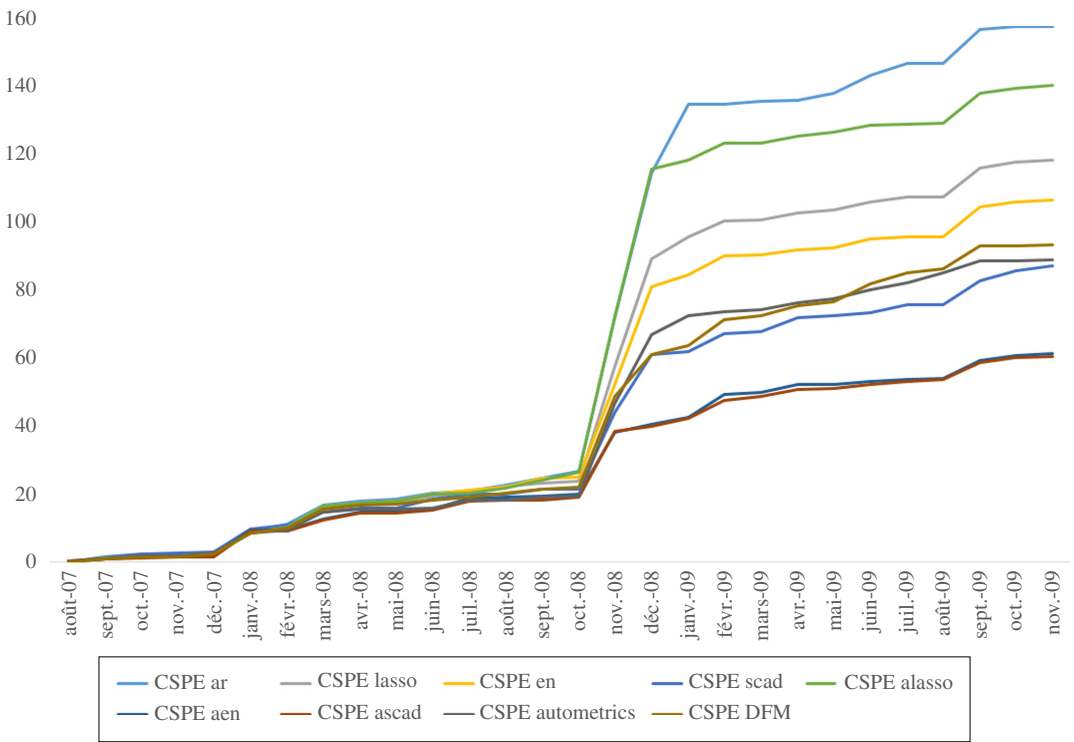


FIGURE 4 Cumulative squared prediction errors for the 2007M8–2009M11 period [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

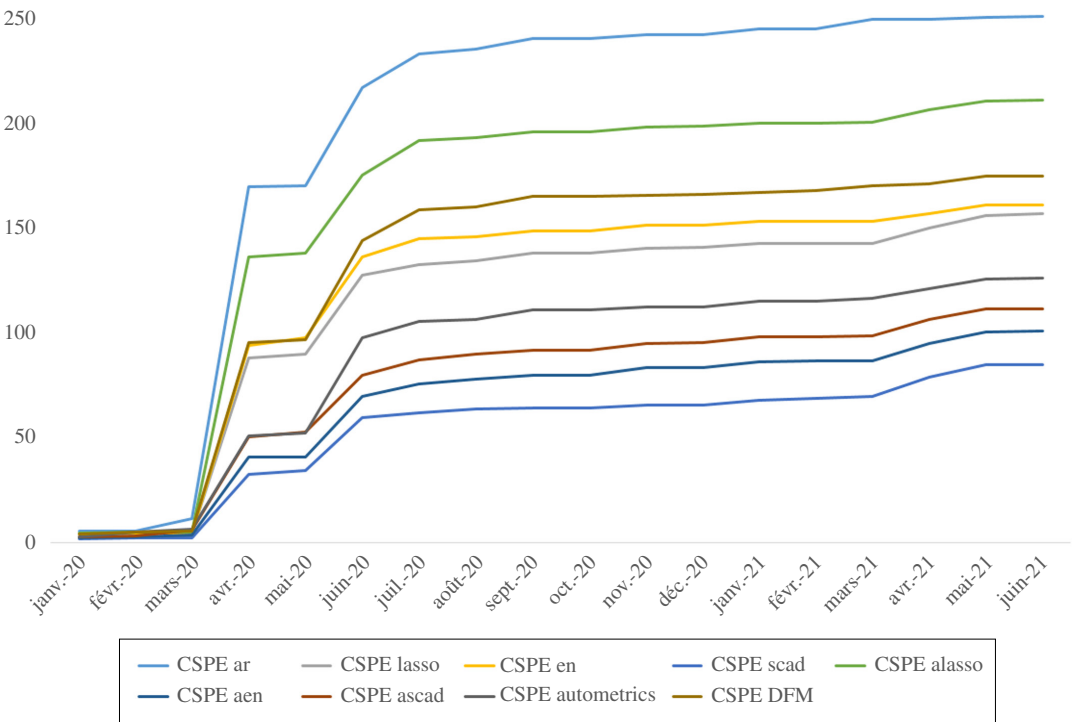


FIGURE 5 Cumulative squared prediction errors for the 2020M1–2021M6 period [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

the DFM is never significantly outperformed and, inversely, it never outperforms the models based on Lasso, aLasso and EN during non-crisis time and those based on SCAD, aSCAD and aEN during crisis times. This result is in contradiction with the results of Burgert and Dees (2009), Guichard and Rusticelli (2011) and Jakaitiene and Dees (2012) who find that the DFMs give better forecasts.<sup>25</sup>

## 6 | CONCLUSION

In this paper, we compared backcasting performance of models based on variable selection to dynamic factor model for backcasting the world trade growth rate with two months ahead. The variable selection models were specified by employing penalised regressions and an econometric approach based on an automatic general-to-specific procedure, using a large data set. A recursive forecast study was carried out to assess the backcasting performance by distinguishing crisis and non-crisis periods. The results showed that, some selection-based models exhibit a good backcasting performance during both periods and are thus relevant for backcasting world trade growth. The accuracy gains from variable selection methods are large during the GFC and COVID-19 crisis where the more accurate backcasts are provided by SCAD, adaptive EN and adaptive SCAD. Small accuracy gains with respects to the DFM (and also to the AR(2) model) are obtained by Lasso, Elastic-Net, adaptive Lasso during the non-crisis episode. Amongst the predictors for backcasting world trade growth it appears that the index of global economic conditions proposed by Baumeister et al. (2020), the PMI indicator on new export orders in manufacturing sector and the MSCI world index are important due to the fact they are selected by all the approaches.

For practitioners, our findings suggest that when factor models and variable selection methods (penalised regressions and GETS approach) give similar performances the latter methods should be preferred since they also allow for interpretability and ability to identify the main drivers of the models.

This paper is focused on backcasting the world trade growth. Further research would be to design the models to forecast world trade to a longer horizon. Another extension could be to nowcast/forecast the world trade growth by restricting to the lags of the regressors instead of using the contemporaneous values. Since we use indicators with mixed frequencies an extension of this work would be to apply the MIDAS version of some penalised techniques such as, for example the Lasso-MIDAS of Babii et al. (2021). For the machine learning approaches, we used penalised regressions. We could also consider other approaches which allow nonlinear specification in future research, such as random forest, boosting methods or support vector machine.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge the two referees for their constructive comments, which have improved the quality of the paper. Any remaining errors are our own. We are also grateful to IHS Markit and Clarksons ResearchTM for providing us PMI data and ClarkSea index, respectively.

<sup>25</sup>This difference could be explained by the fact that (i) Guichard and Rusticelli (2011) forecast OECD world trade growth and not CPB world trade; (ii) Burgert and Dees (2009) and Jakaitiene and Dees (2012) forecast CPB world trade but not in backcasting; and (iii) the sample period of analysis is not the same with these three studies.

## ORCID

Olivier Darné  <https://orcid.org/0000-0003-2796-4873>

## REFERENCES

- Babii, A., Ghysels, E., & Striaukas, J. (2021). Machine learning time series regressions with an application to nowcasting. *Journal of Business & Economic Statistics*, forthcoming. <https://doi.org/10.1080/0735015.2021.1899933>
- Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. *Quarterly Journal of Economics*, 131, 1593–1636. <https://doi.org/10.1093/qje/qjw024>
- Barhoumi, K., Darné, O., & Ferrara, L. (2013). Dynamic factor models: A review of the literature. *Journal of Business Cycle Measurement and Analysis*, 8, 73–107. <https://doi.org/10.2139/ssrn.2291459>
- Barhoumi, K., Darné, O., & Ferrara, L. (2016). A World Trade Leading Index (WTLI). *Economics Letters*, 146, 111–115.
- Baumeister, C., Korobilis, D., & Lee, T. K. (2020). Energy markets and global economic conditions. *The Review of Economics and Statistics*, forthcoming. [https://doi.org/10.1162/rest\\_a\\_00977](https://doi.org/10.1162/rest_a_00977)
- Burgert, M., & Dees, S. (2009). Forecasting world trade: Direct versus bottom-up approaches. *Open Economic Review*, 20, 385–402. <https://doi.org/10.1007/s11079-007-9068-y>
- Caldara, D., & Iacoviello, M. (2018). *Measuring geopolitical risk*. International Finance Discussion Papers No 1222, Board of Governors of the Federal Reserve System.
- Caldara, D., Iacoviello, M., Molligo, P., Prestipino, A., & Raffo, A. (2020). The economic effects of trade policy uncertainty. *Journal of Monetary Economics*, 109, 38–59. <https://doi.org/10.1016/j.jmoneco.2019.11.002>
- Campos, J., Hendry, D. F., & Ericsson, N. R. (2005). *General-to-specific modeling* (Vol. 1 and 2). Edward Elgar Publishing.
- Castle, J. L. (2005). Evaluating PcGets and RETINA as automatic model selection algorithms. *Oxford Bulletin of Economics and Statistics*, 67, 837–880. <https://doi.org/10.1111/j.1468-0084.2005.00143.x>
- Doornik, J. A. (2009). Autometrics. In J. L. Castle, & N. Shephard (Eds.), *The Methodology and Practice of Econometrics* (pp. 88–121). Oxford: Oxford University Press.
- Doornik, J. A., & Hendry, D. F. (2018). *Empirical econometric modelling – PcGive TM 15* (Vol. 1). Timberlake.
- Doz, C., Giannone, D., & Reichlin, L. (2011). A two-step estimator for large approximate dynamic factor models based on Kalman filtering. *Journal of Econometrics*, 164, 188–205. <https://doi.org/10.1016/j.jeconom.2011.02.012>
- Doz, C., Giannone, D., & Reichlin, L. (2012). A quasi maximum likelihood approach for large approximate dynamic factor models. *Review of Economics and Statistics*, 94, 1014–1024. [https://doi.org/10.1162/REST\\_a\\_00225](https://doi.org/10.1162/REST_a_00225)
- Epprecht, C., Guégan, D., Veiga, Á., & Correa da Rosa, J. (2021). Variable selection and forecasting via automated methods for linear models: LASSO/adaLASSO and Autometrics. *Communications in Statistics – Simulation and Computation*, 50, 103–122. <https://doi.org/10.1080/03610918.2018.1554104>
- Fan, J., & Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, 96, 1348–1360. <https://doi.org/10.1198/016214501753382273>
- Fan, J., & Peng, H. (2004). Nonconcave penalized likelihood with a diverging number of parameters. *Annals of Statistics*, 32, 928–961. <https://doi.org/10.1214/009053604000000256>
- Forni, M., Hallin, M., Lippi, M., & Reichlin, L. (2005). The generalized dynamic factor model: One-sided estimation and forecasting. *Journal of the American Statistical Association*, 100, 830–840. <https://doi.org/10.1198/016214504000002050>
- Guichard, S., & Rusticelli, E. (2011). *A dynamic factor model for world trade growth*. Working papers No 874, Economics Department, OECD.
- Hansen, P. R. (2005). A test for superior predictive ability. *Journal of Business & Economic Statistics*, 23, 365–380. <https://doi.org/10.1198/073500105000000063>
- Hansen, P. R., Lunde, A., & Nason, J. M. (2011). Model confidence sets for forecasting models. *Econometrica*, 79, 453–497.
- Jakaitiene, A., & Dees, S. (2012). Forecasting the world economy in the short term. *The World Economy*, 35(3), 331–350. <https://doi.org/10.1111/j.1467-9701.2011.01433.x>

- Kim, H. H., & Swanson, N. R. (2014). Forecasting financial and macroeconomic variables using data reduction methods: New empirical evidence. *Journal of Econometrics*, 178, 352–367. <https://doi.org/10.1016/j.jeconom.2013.08.033>
- Kim, H. H., & Swanson, N. R. (2018). Mining big data using parsimonious factor, machine learning, variable selection and shrinkage methods. *International Journal of Forecasting*, 34, 339–354. <https://doi.org/10.1016/j.ijforecast.2016.02.012>
- Li, J., & Chen, W. (2014). Forecasting macroeconomic time series: LASSO-based approaches and their forecast combinations with dynamic factor models. *International Journal of Forecasting*, 30, 996–1015. <https://doi.org/10.1016/j.ijforecast.2014.03.016>
- Martínez-Martín, J., & Rusticelli, E. (2021). Keeping track of global trade in real time. *International Journal of Forecasting*, 37, 224–236. <https://doi.org/10.1016/j.ijforecast.2020.04.005>
- Smeekes, S., & Wijler, E. (2018). Macroeconomic forecasting using penalized regression methods. *International Journal of Forecasting*, 34, 408–430. <https://doi.org/10.1016/j.ijforecast.2018.01.001>
- Stock, J. H., & Watson, M. W. (2002). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20, 147–162. <https://doi.org/10.1198/073500102317351921>
- Stock, J. H., & Watson, M. W. (2015). Factor models for macroeconomics. In J. B. Taylor, & H. Uhlig (Eds.), *Handbook of macroeconomics* (Vol. 2, pp. 415–525). North Holland.
- Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society: Series B*, 58, 267–288.
- van Welzenis, G., & Suyker, W. (2005). *Explanatory note on the CPB world trade series*. CPB Memorandum No 116, CPB Netherlands Bureau for Economic Policy Analysis.
- World Trade Organization (2016). *World trade outlook indicator (WTOI)*. WTO Methodological note. <https://www.wto.org>
- Xiao, N., & Xu, Q.-S. (2015). Multi-step adaptive elastic-net: Reducing false positives in high-dimensional variable selection. *Journal of Statistical Computation and Simulation*, 85, 1–11. <https://doi.org/10.1080/00949655.2015.1016944>
- Zou, H. (2006). The adaptive Lasso and its oracle properties. *Journal of the American Statistical Association*, 101, 1418–1429. <https://doi.org/10.1198/016214506000000735>
- Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society, Series B*, 67, 301–320. <https://doi.org/10.1111/j.1467-9868.2005.00503.x>
- Zou, H., & Zhang, H. H. (2009). On the adaptive elastic-net with a diverging number of parameters. *Annals of Statistics*, 37, 1733–1751. <https://doi.org/10.1214/08-AOS625>

**How to cite this article:** Charles, A., & Darné, O. (2022). Backcasting world trade growth using data reduction methods. *The World Economy*, 45, 3169–3191. <https://doi.org/10.1111/twec.13274>

## APPENDIX A

## Data

TABLE A1 Information sources of the data

| Name                                 | Start date | Leads | Transform. | Information         | Source                        |
|--------------------------------------|------------|-------|------------|---------------------|-------------------------------|
| World trade                          | 2000       | 0     | 2          | World trade         | CPB                           |
| Global Economic Conditions indicator | 1973       | 2     | 0          | Economic conditions | Baumeister et al. (2020)      |
| Baltic Dry Index                     | 1985       | 2     | 2          | Maritime freight    | Datastream                    |
| ClarkSea index                       | 1990       | 2     | 2          | Maritime freight    | Clarksons Research™           |
| Brent crude oil prices               | 1957       | 2     | 2          | Commodity market    | EIA                           |
| Commodity Research Bureau index      | 1991       | 2     | 2          | Commodity market    | CRB                           |
| S&P GSCI commodity index             | 1973       | 2     | 2          | Commodity market    | Datastream                    |
| World MSCI index                     | 1973       | 2     | 2          | Stock market        | Datastream                    |
| USD nominal effective exchange rate  | 1963       | 2     | 2          | Exchange rate       | BIS                           |
| US Treasury yield spread             | 1984       | 2     | 0          | Bond market         | FRED                          |
| Trade policy uncertainty             | 1960       | 2     | 1          | Uncertainty         | Caldara et al. (2020)         |
| Geopolitical risk                    | 1985       | 2     | 1          | Uncertainty         | Caldara and Iacoviello (2018) |
| US trade policy uncertainty          | 1985       | 2     | 1          | Uncertainty         | Baker et al. (2016)           |
| Global PMI composite input prices    | 1998       | 2     | 0          | Survey              | IHS Markit                    |
| Global PMI composite new orders      | 1998       | 2     | 0          | Survey              | IHS Markit                    |
| Global PMI composite output          | 1998       | 2     | 0          | Survey              | IHS Markit                    |
| Global PMI manuf input prices        | 1998       | 2     | 0          | Survey              | IHS Markit                    |
| Global PMI manuf new orders          | 1998       | 2     | 0          | Survey              | IHS Markit                    |

TABLE A1 (Continued)

| Name                               | Start date | Leads | Transform. | Information | Source     |
|------------------------------------|------------|-------|------------|-------------|------------|
| Global PMI manuf new export orders | 1998       | 2     | 0          | Survey      | IHS Markit |
| Global PMI manuf output            | 1998       | 2     | 0          | Survey      | IHS Markit |
| Global PMI manuf purchase stocks   | 1998       | 2     | 0          | Survey      | IHS Markit |
| Global PMI manuf index             | 1998       | 2     | 0          | Survey      | IHS Markit |
| IFO climate                        | 1991       | 2     | 0          | Survey      | IFO        |
| IFO expectation                    | 1991       | 2     | 0          | Survey      | IFO        |

*Note:* Leads denotes the lead of the predictors on the publication of the monthly index of world trade by the CPB. Transform. denotes the transformation applied on the variables with the following codes: 0—no transformation; 1—first difference; 2—first difference of logarithm.