

Stamer, Vincent

**Conference Paper**

## Thinking Outside the Container: A Sparse Partial Least Squares Approach to Forecasting Trade Flows

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2022: Big Data in Economics

**Provided in Cooperation with:**

Verein für Socialpolitik / German Economic Association

*Suggested Citation:* Stamer, Vincent (2022) : Thinking Outside the Container: A Sparse Partial Least Squares Approach to Forecasting Trade Flows, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2022: Big Data in Economics, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/264096>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Thinking Outside the Container: A Sparse Partial Least Squares Approach to Forecasting Trade Flows\*

Vincent Stamer<sup>†</sup>

February 28, 2022

## Abstract

Global container ship movements may reliably predict trade flows both in normal times and in times of economic crisis. This paper outlines the methods including a sparse partial least squares model to construct forecasts for trade flows from several million container ship positions per year. Out-of-sample exercises show that constructed forecasts for 76 countries and regions outperform benchmark models and improve model combinations. This holds true for unilateral and bilateral trade flows of developed and developing countries. Forecasts of imports may precede official statistics by several months and may facilitate quantification of supply chain disruptions and trade wars.

**Keywords:** Forecasting, Trade, Forecasting, Factor Model, Partial Least Squares, Container Shipping

**JEL classification:** F47; C53; E66

---

\*The author greatly appreciates the research assistance of Steffen Gans, Sven-Niklas Waldow and Falk Wendorff. Gratitude is owed to Gabriel Felbermayr and Jens Boysen-Hogrefe for continued support, as well as participants of the ETSG conference, IfW-lunchtime seminar, CAU econometrics seminar, WIFO lunchtime seminar, IFW-staff, Serkan Arslanalp, Diego Alejandro Cerdeiro, Nils Jannsen, Stefan Kooths, Andras Komaromi, Saskia Meuchelböck, Mariia Okuneva and Yang Liu for valuable advice.

<sup>†</sup>Kiel University, Kiel Institute for the World Economy; Wilhelm-Seelig-Platz 1, 24118 Kiel; E-mail: vincent.stamer@ifw-kiel.de

# 1 Introduction

As the coronavirus spread across the globe, it disrupted the trade of goods in its wake. In April 2020 for instance, Germany and France experienced year-over-year export drops of 33.6 % and 44.5 %, respectively (WTO, 2020). Even two years after the first outbreak, supply chain disruptions continue to limit the economic recovery. Celasun et al. (2022) find that the euro area's GDP in 2021 would have been 2% higher in the absence of supply shortages. To mitigate the uncertainty associated with plummeting trade volumes and supply shortages, decision makers in governments and corporations require the most current information. Yet, even in developed countries, statistical offices publish unilateral trade data with a lag of several weeks and bilateral data often with a lag of months.

This study applies a sparse partial least squares (PLS) algorithm and other machine learning techniques on daily container ship movements to fill this gap. Accounting for almost half of global trade by value, container ships moved more than 8 trillion USD worth of goods across the seas in 2020 (UNCTAD, 2021). At the same time, positional data of container ships is available at a high frequency in a matter of hours. This study first spatially aggregates over two million container ship positions per year to around 1,200 predictor time series using a k-means clustering algorithm. Using these predictor time series, a sparse partial least squares algorithm constructs models of unilateral and bilateral trade flows. This allows forecasts of trade flows one month ahead, which therefore predate official statistics by several months. Out of sample tests show that the forecasts derived from the container ship positions at the very least contribute new information to benchmark models. For the majority of countries and trade flows, the partial least squares model significantly outperforms benchmark models. Hence, this paper addresses not only the question how to derive forecasts from highly disaggregated geocoded data, but also how well the forecasts perform against benchmarks.

As container ship movements can be spatially aggregated to around 1,200 predictor time series, dimensionality reduction plays a key role in the analysis. This paper follows a growing literature using factor models to reduce dimensionality (see Stock and Watson

2002; Bai and Ng 2004; Eickmeier and Ng 2011; Jurado et al. 2015). For instance, Eickmeier and Ng (2011) find that PLS is well suited to integrate rich international indicators to model national GDP. At the same time, Boivin and Ng (2006) have warned that a higher number of explanatory series in the cross-section may not be desirable in factor models if idiosyncratic errors are cross-correlated. Similarly, Camacho and Perez-Quiros (2010) successfully forecast euro area GDP using a small-scale factor model that dramatically limits the number of predictors. As the number of predictors in this study is high even for factor models, explanatory time series are pre-selected using the variable influence on projection (VIP) as used by Wold et al. (2001) and a custom algorithm taking into account the magnitude and geographic proximity of container ship ports. While selection techniques differ, the methodologically most closely related study is Fuentes et al. (2015). In their PLS model of inflation, the authors refer to their predictor selection procedure as sparse partial least squares (SPLS). This paper contributes to this strand of econometric methodology by introducing to economics a different selection method in the PLS model, as well as by outlining aggregation methods to geocoded data.

Above studies typically integrate explanatory time series from a variety of sources for a full forecast of one dependent time series. This work flips this approach: The aim is to provide forecasts for many trade flows from one source of information, the container shipping network. National practitioners can, then, integrate these forecasts as leading indicators in larger, country-specific forecast models. Hence, this work tests forecasts from the shipping data in a pseudo out-of-sample, expanding window exercise and compares them against benchmark ARX models that include autoregression terms and competing leading indicators. A Diebold-Mariano test (Diebold and Mariano, 2002) and a simple encompassing test following Davidson et al. (1993) show that the forecasts from the shipping network add new information to existing models and frequently outperform them. Results hold both for the years 2015 to 2019 and for the first half of 2020, which saw significant trade disruptions due to the COVID-19 pandemic. In fact, the new indicator outperforms autoregression models much more strongly in times of crisis. Additionally, forecasts of trade flows can be derived for any desired forecasting horizon,  $h$ , in a direct-step forecast-

ing fashion and at any point of the month. This paper contributes to forecasting work by testing the new forecasts as leading indicators.

The Automatic Identification System (AIS), which records ship positions and other information based on ships' radio signals, represents the primary data source for this study. For other applications of AIS data in economics, see, for instance, Brancaccio et al. (2017), Heiland et al. (2019) and Wong and Ziv (2020). Two proof-of-concept studies of the IMF also use AIS data and are very related to this work: Arslanalp et al. (2019) derive a trade indicator from the activity of two Maltesian ports and compare it to Malta's official trade data. Cerdeiro et al. (2020) develop methods to calculate port calls (stops of a ship at a port to load or unload cargo) from ship positions. They also show that the net-draught change of ships entering and departing from ports tends to correlate positively with unilateral imports and exports. This paper serves as a natural extension of Cerdeiro et al. (2020) as it shows how shipping data can be used and tested in forecasting. Further, this paper contributes to the economic analysis of shipping data through the introduction of a PLS model to this field. Factor models exploit any source of correlation between ship movements and a country's trade flows that arise through common trends of countries, input-output linkages across countries, as well as substitution and complementary relationships that transcend borders. This paper, therefore, introduces advanced methods that are well suited for a wide variety of other data sources such as mobility, production and retail data.

The following chapter describes the data used and the necessary steps to make AIS data interpretable to a factor model. The section on methodology describes the partial least squares model and the required time series tests. The results chapter performs these tests and presents results including comparisons to benchmark models. The section also highlights the strength of the new indicator in times of crises. The discussion section deals with the dependency of the results on the methodology and trade data used. For instance, a robustness test uses nominal trade data instead of real trade data as dependent time series. A second test shows that results hold even if only half the months shipping data is used for forecasting. The last chapter concludes and gives an outlook for work ahead.

## 2 Data

### 2.1 Data Collection of Predictor Time Series

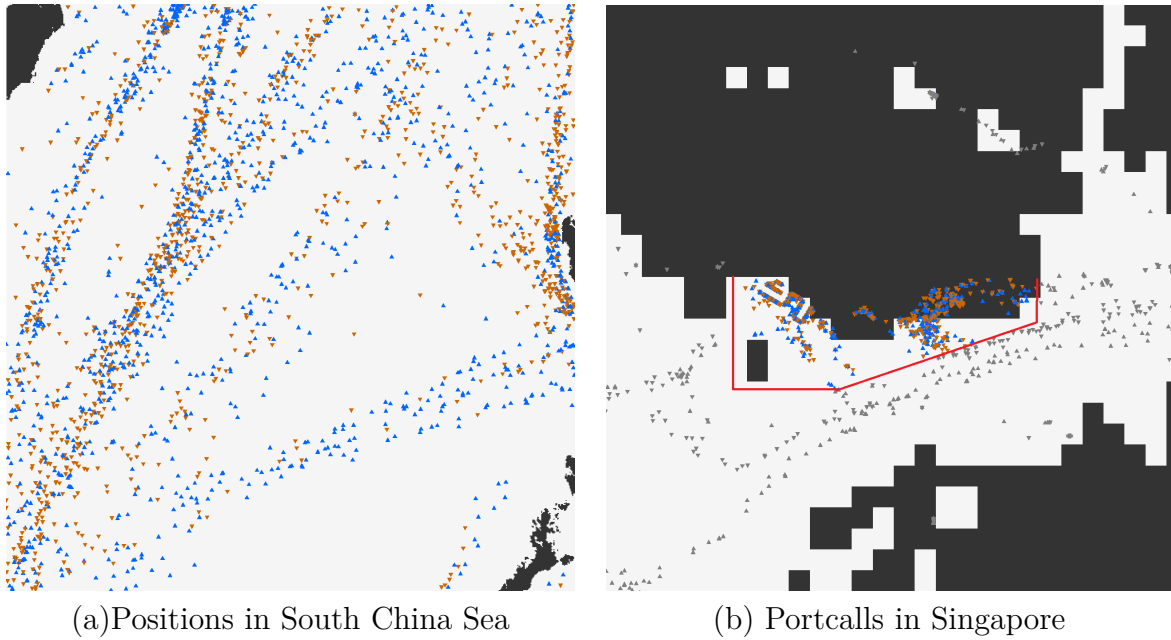
The main data source is the Automatic Identification System (AIS), a radio system to monitor ship movements and avoid collisions at sea. Ships broadcast radio signals every few seconds containing ship identifiers, course, speed, GPS position and draught (vertical distance between the waterline and the bottom of the ship’s hull). The data provider Fleetmon.com collects the signals from terrestrial receiver stations and satellites. The agency also determines whenever a ship stops at a port to load or unload cargo, a so called port call. Fleetmon.com provides this project with a proprietary dataset of one AIS position per day and the exhaustive set of all port calls for all container ships worldwide for the time span between January 2015 to October 2021. This data is combined with ship information from MarineTraffic such as size, minimum draught (“ballast draught”) and capacity measured in Twenty-Foot-Equivalent units (TEU).

For the AIS data to be useful for economic forecasting, one must further organize the data into time series of independent variables. To organize the positions at sea, the first step partitions the world between 50°latitude South and 70°latitude North into areas of 10°latitude and 10°longitude. The 100 areas with the greatest number of container ship positions are kept while the rest is discarded to avoid overfitting the factor models at a later stage. Panel (a) of Figure 1 shows one such area in the South China Sea and container ship positions recorded in the month of July 2020. Orange triangles represent ships moving on a Southern course and blue triangles represent ships on a Northern course. Determining two predominant directions follows the intuition that ships going in the opposite direction likely carry cargo destined for different countries. A simple k-means clustering algorithm (Hartigan and Wong, 1979) assigns ships to a predominant course once the course in degrees is mapped into two-dimensional space using the sine and cosine functions.<sup>1</sup>

---

<sup>1</sup>The course in the AIS message is given in degrees from zero to 360 and standard clustering algorithms fail to match the values zero and 360 to the same course without projection into two-dimensional space.

Figure 1: Illustration of ship data aggregation



In addition to the positions, port calls provide the information when ships enter into and depart from a port such as the port of Singapore in panel (b) of Figure 1. In this largely oversimplified illustration, orange triangles signify ships departing from the port while blue triangles show ships that have just entered the port. The information such as the draught is recorded as the ships cross the imaginary red line, a step calculated by the data provider Fleetmon.com.

All ships, regardless of whether they enter or depart a port area or whether they cruise in one of the remaining areas at sea, are assigned their approximate cargo load. This study uses the current draught of a ship to calculate the load measured in TEU. After all, a ship that approaches a port with a draught close to its minimum value likely carries less containers than a ship cruising with a draught near its maximum, *ceteris paribus*. Hence, the current draught is normalized by the potential range of draughts and multiplied by the maximum TEU capacity of the ship to approximate the number of containers the ship currently carries.

$$TEUload_{it} = TEU_{i,max} \times \frac{draught_{it} - draught_{i,min}}{draught_{i,max} - draught_{i,min}} \quad (1)$$

Formula (1) translates this into an equation for the TEU load of ship  $i$  at time  $t$ . It also highlights a subtle difference between this approach and other studies: While Cerdeiro et al. (2020) subtract the departure load from the arrival load to calculate a net gain of the ship's cargo during a port call, both the gross load of the ship while entering and while departing are kept as two separate time series. This approach has a key advantage in the following situation: Whenever a ship unloads containers in a port for import and loads containers for export, the draught of the ship entering and exiting the port may coincidentally be identical. Hence, only the two gross values of entering and exiting account for import and export values simultaneously. In theory, this approach may overstate the importance of transshipment hubs, i.e. ports such as the port of Singapore where containers are primarily transferred between ships. In this case, a fully loaded container ship does not necessarily reflect imports of the state of Singapore. In practice, however, the PLS algorithm automatically accounts for transshipment effects precisely because there will not be a correlation between ship movements and trade statistics. One additional reason to use the gross load of ships at port call events is that the positions at sea only have a gross TEU load value.

Lastly, summing the monthly TEU loads of all ships by predominant course yields two observational time series per area. Monthly time series per port sum the TEU load of ships entering the port and of ships departing the port. Again, this yields two series per port. In total, the 100 areas and 500 ports generate 1200 time series serving as independent variables. To avoid ruggedness and increase flexibility of the setup, the cutoff day for summing monthly data does not have to coincide with the last day of the calendar month. For instance, setting the cutoff day to be the 15th of a month implies summing the data for the 30 days leading up to the 15th of a month. This also generates time series of 12 time periods per year and allows the researcher to produce nowcasts for the current month shortly after the 15th. Except when stated otherwise, however, the cutoff date in this study is set to be the 30th of the month.



## 2.2 Processing of Predictor Time Series

Assumptions of the PLS model as outlined in the section on methodology necessitate several processing steps of the shipping derived, independent time series. The first step is to reduce further the number of predictor time series. To that end, the ports are ranked by average monthly TEU load and the median volume by port,  $M$ , is calculated. Iteratively the port with the lowest monthly TEU load is added to the closest port by geographic proximity. This is repeated until no port displays a volume below the original median,  $M$ . This reduces the number of time series to approximately 750 and significantly reduces the variance of the smallest ports' volume remaining in the dataset. The second step derives stationary, month-over-month growth rates from above time series. Using the X-13 procedure of the US Census Bureau (2017) on both predictor and target time series seasonally adjusts the data and reduces cross-correlation in the error terms. This is because clusters of ports in close geographic proximity may be subject to similar seasonal trends through weather conditions, for instance.

## 2.3 Other Data

Data sources for dependent time series include CPB (2021) for price adjusted, unilateral trade flows, the WTO (2020) for nominal, unilateral trade flows and IMF (2021) for monthly bilateral trade. European economic sentiment indicators (ESI) and industrial confidence indicators (COF) originate from the European Commission (2020), US industrial production (US IP) from the Federal Reserve Bank of St. Louis (2020), both German export expectations (Grimme and Wohlrabe, 2014) and import and export climate (Grimme et al., 2018) from the ifo institute (Ifo institute, 2021), German truck toll index, order entry and order stock from abroad originate from DESTATIS (2021) and a port throughput index (Döhrn and Maatsch, 2012) comes from RWI/ISL (2020). The benchmark models described below exclusively use the seasonally adjusted versions of the indicators and utilize a time series as long as possible.

## 3 Methodology

### 3.1 Partial Least Squares

Wold (1975) introduced the partial least squares (PLS) model as early as 1975 and several handbooks describe the procedure in detail (see, for instance, Garthwaite 1994; Haenlein and Kaplan 2004; Esposito Vinzi et al. 2010; Lohmöller 2013). Factor models such as the principle component regression (PCR) and PLS have gained popularity because they reduce dimensionality in the span of the predictor time series particularly in the presence of strong colinearity among predictors (Wold et al., 2001). As the number of predictors in the cross-section ( $N$ ) and the number of observations of the time series ( $T$ ) tend to infinity, PLS makes consistent estimates of the factor space under the following assumptions: Estimation errors are stationary, components have non-trivial loadings and errors have at most weak cross-correlation (Stock and Watson, 2002; Bai and Ng, 2002). Chun and Keleş (2010) relax the limiting assumption of  $N$  and  $T$  going to infinity. The authors establish that consistent estimates are achieved if and only if  $N$  grows more slowly than  $T$ . Similarly, Rönkkö (2014) caution that as  $N$  grows the likelihood of chance correlations rises.

This has two implications for this study: First, as the threat of inconsistent estimates and overfitted models looms, evaluating the forecast quality in an out-of-sample study gains importance. While this does not remedy overfitting, it measures the resulting imprecision of the model. Second, the number of predictor time series  $N$  must be reduced below the number of observations  $T$  before running PLS. This is partly achieved in the aggregation of ports as outlined in the section on data. Additionally, predictor time series are pre-selected using the metric variable influence on projection (VIP), which has gained traction in chemometrics (Tran et al., 2014; Galindo-Prieto et al., 2014; Cocchi et al., 2018).

The following summarizes the partial least squares process: Factor models such as PLS rely on the idea that explanatory time series  $X$  are expressions of a small number of latent factors  $F$  and an idiosyncratic error term  $\varepsilon$ . Factor models approximate these

latent factors and use them as components in a model for the dependent time series  $y$ .

$$X_t = \Lambda F_t + \varepsilon_t \quad (2)$$

$$y_{t+h} = \alpha + \beta \hat{F}_t + \eta_{t+h} \quad (3)$$

where  $X_t$  is an  $N$ -length vector of predictors observed at time  $t$  and  $F_t$  refers to the  $k$ -dimensional latent variables, or components, vector. The PLS-algorithm estimates the  $N \times k$  matrix of so called factor loadings,  $\Lambda$ . Ordinary least squares provides estimates of the  $1 \times k$  vector of coefficients,  $\beta$ . Error terms are shown by vector  $\varepsilon_t$ . The time subscript,  $t$ , is dropped for the following description of components for ease of notation.

An individual component  $F_k$ , *for*  $k = 1, \dots, K$ , can be aggregated from the  $J$  centered explanatory variables  $V_{jk}$ , *for*  $j = 1, \dots, J$ , using

$$F_k = \sum_{j=1}^J w_{jk} \phi_{jk} V_{jk} \quad (4)$$

$$V_{jk} = X_{jk} - \bar{\mathbf{x}}_{jk} \quad (5)$$

One derives the component and explanatory variable specific correlation coefficients  $\phi_{jk}$  from individually regressing the centered dependent variable  $U_k$  sequentially on each of the explanatory variables directly:

$$U_k = Y_k - \bar{\mathbf{y}}_k \quad (6)$$

$$U_k = \phi_{jk} V_{jk} \quad (7)$$

Furthermore, the algorithm calculates the components iteratively, so that the first component is constructed from the original dependent and independent time series. The second component uses only the residuals of the regression of  $U_{k=1}$  on  $F_{k=1}$  as the new dependent variable  $U_{k=2}$  and uses the residuals of the regressions of  $V_{j,k=1}$  on  $F_{j,k=1}$  as the new explanatory variables  $V_{j,k=2}$ .

The choice of weights  $w_{jk}$  depends on the specific PLS algorithm. Fuentes et al. (2015),

for instance, set the weights vector to maximize the covariance between dependent and the independent time series:

$$w = \arg \max_w w'V'UU'Vw \quad (8)$$

Note that in contrast to dynamic applications of PLS such as (Eickmeier and Ng, 2011), this study does not take into account lags of the explanatory variables. In this study, lags of shipping time series are mostly redundant as ships move across the oceans over time. The loads carried by freight ships departing China for the Pacific Ocean today, for example, will be observed nearing the American West Coast in several weeks. Hence, observations of areas near the American West Coast are highly correlated with observations of areas near Asia several weeks ago.

The variable influence on projection (VIP) statistic can be calculated for predictor time series  $j$  as such:

$$VIP_j = \sqrt{K \frac{\sum_{k=1}^K SSY_k \left( \frac{w_{kj}}{\|w_k\|} \right)^2}{SSY_{total}}} \quad (9)$$

where  $SSY_k$  is the sums of squares explained by individual component  $k$  and  $SSY_{total}$  is the sums of squares explained by the entire factor model. The sums of squares explained by components are multiplied by the weights of the individual predictor variables to reflect the predictor's influence on the dependent time series through the various components. VIP identifies the 30 most important explanatory variables to be kept for the actual fitting of the model.

Next, cross-validation determines the optimal number of components  $K$  for the specific dependent time series in line with previous studies (Wold et al., 1984, 2001). The maximum number of components is constrained to 10 to avoid overfitting, although rarely more than three components are optimal.

### 3.2 Forecast evaluation

To evaluate the forecast quality of the algorithm above, PLS calculates expanding window out-of-sample forecasts for the second half of the sample period. Let the forecast horizon,  $h$ , - specific prediction for the seasonally adjusted monthly growth rate of the dependent variable be  $\widehat{y_{t+h}^{PLS}}$ . Both a direct comparison to benchmark models using the Diebold-Mariano test and an encompassing test assess the forecast quality. The expanding window approach calculates out-of-sample predictions beginning in May 2018 using half the available data.

The PLS forecasts are compared to a naive model set to the mean growth rate and an ARX model with distributed lags up until  $p = 3$ . The Akaike information criterion determines the number of lags and indicators used. As the PLS inherently forecasts in a direct-step fashion, the ARX fitting follows accordingly:

$$y_{t+h} = \bar{y} \tag{10}$$

$$y_{t+h} = \alpha + \sum_{i=1}^p \beta_i y_{t-i} + \sum_{i=0}^p \sum_{j=0}^n \gamma_{ji} \text{Indicator}_{j,t-i} + \varepsilon_{t+h} \tag{11}$$

where  $y_{t+h}$  denotes the month-over-month growth rate of the adjusted time series from time period  $t + h - 1$  to  $t + h$ . The intercept  $\alpha$  and coefficients  $\beta_i$  and  $\gamma_{ji}$  are estimated in an expanding window method in the same fashion as the actual shipping indicator. The selection of indicators  $j$  is country specific, but always includes the US industrial production. For all countries covered by the European survey program, the country specific economic sentiment indicator and manufacturing confidence index are used. For Germany, leading indicators additionally include export and import specific indicators such as ifo export expectations, ifo export and import climate, manufacturing order entries, as well as port throughput indicators and a truck activity index. The fitting period for benchmark ARX indicator models dates back as long as possible limited only by the availability of the indicator or the dependent time series.

A standard root mean squared forecast error (RMSFE) calculation evaluates the difference between the out-of-sample forecasts and the actual value specific to the forecast horizon:

$$RMSFE_h = \sqrt{\frac{1}{T} \sum_{t=1}^T (FE_{t+h})^2} \quad (12)$$

This quantifies the accuracy of the models. A relative RMSFE (rRMSFE) divides the forecast error of the ARX-PLS model by the benchmark error.

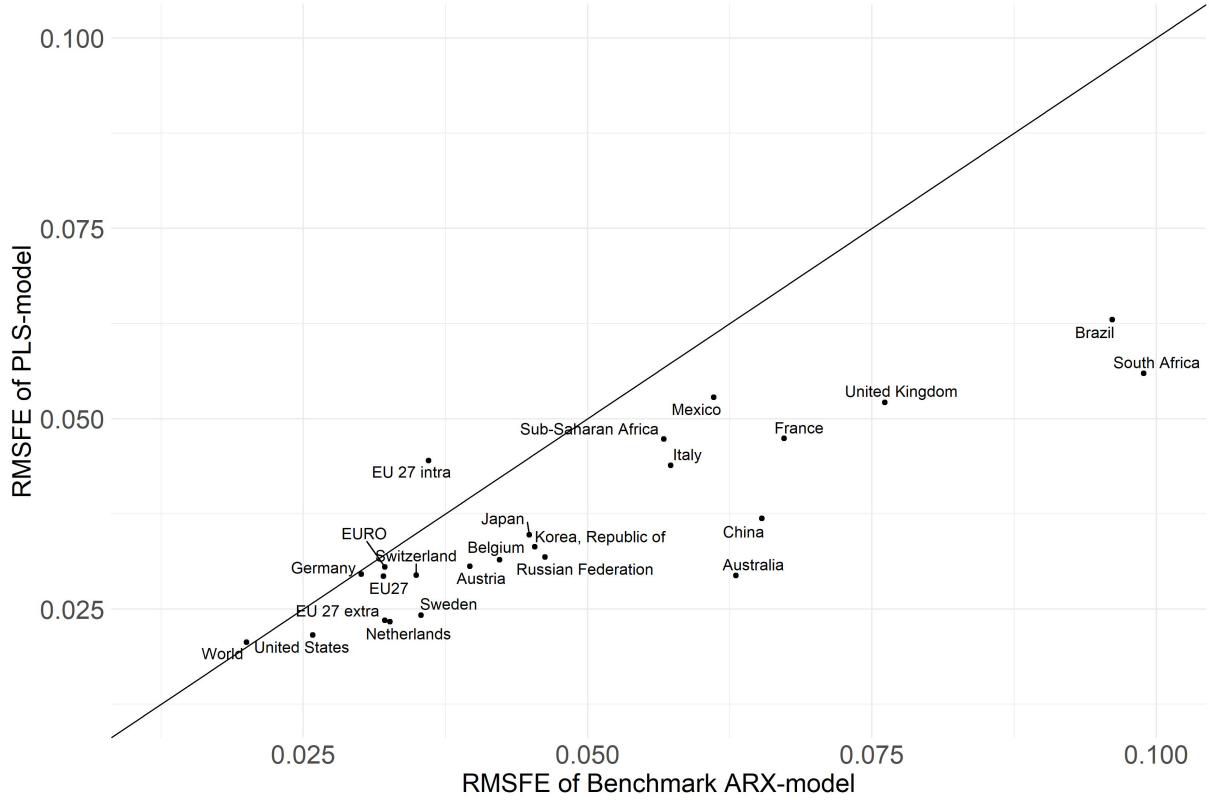
$$relative\ RMSFE_h = \frac{RMSFE_h^{PLS}}{RMSFE_h^{Other}} \quad (13)$$

Hence, a value below 1 for the relative RMSFE indicates lower forecast errors for the new shipping model. Note, that the RMSFE of the naive model is identical to the standard deviation of the time series. The relative RMSFE of the PLS against the naive model, therefore, measures the forecast error against the standard deviation as advocated by Breitung and Knüppel (2021). Lastly, the Diebold-Mariano test (Diebold and Mariano, 2002) establishes the significance of the difference between the PLS and the benchmark models.

To sharpen the intuition for the relative RMSFE values, Figure 2 previews results of forecasting European imports. In particular, the figure compares the RMSFE of the PLS forecast on the y-axis against the RMSFE of the ARX benchmark forecast on the x-axis. A point below the 45°line shows that the PLS model produces lower forecast errors than the benchmark model, i.e. the relative RMSFE is below 1. The figure also illustrates that the lower the errors, the more difficult it is to ascertain which model performs better. For the large unilateral imports of Germany, the EU and the euro area, the autoregressive terms of the models capture a major portion of the import variation.

The Diebold-Mariano test may show that for some target time series, the benchmark model significantly outperforms the PLS model. The PLS model may nevertheless hold predictive power for the target time series. To test this, this paper constructs an en-

Figure 2: Comparison of forecast errors of PLS shipping model (y-axis) against ARX benchmark (x-axis),  $h=0$



compassing test following Davidson et al. (1993). First, a model combination, called an encompassing model, combines regressors both from the benchmark model and the estimate of the PLS model.

$$y_{t+h} = \alpha + \sum_{i=1}^p \beta_i y_{t-i} + \sum_{i=0}^p \sum_{j=0}^n \gamma_{ji} \text{Indicator}_{j,t-i} + \delta \widehat{y_{t+h}^{PLS}} + \varepsilon_{t+h} \quad (14)$$

Let the predictions by the encompassing model be denoted by  $\widehat{y_{t+h}^{Encomp}}$  and those of the benchmark model by  $\widehat{y_{t+h}^{ARX-Bench}}$ . The encompassing test proceeds by calculating the Wald statistic for the two null hypotheses

$$H_0^{PLS} : \lambda_1 = 0 \quad (15)$$

$$H_0^{ARX-Bench} : \lambda_2 = 0 \quad (16)$$

where  $\lambda_1$  and  $\lambda_2$  are implicitly defined as

$$\widehat{y_{t+h}^{Encomp}} = \alpha + \lambda_1 \widehat{y_{t+h}^{PLS}} + \varepsilon_{t+h} \quad (17)$$

$$\widehat{y_{t+h}^{Encomp}} = \alpha + \lambda_2 \widehat{y_{t+h}^{ARX-Bench}} + \varepsilon_{t+h} \quad (18)$$

If the Wald test rejects the null hypothesis  $H_0^{PLS}$ , the PLS is not being encompassed and does hold predictive power beyond competing models. Similarly, if the null hypothesis  $H_0^{ARX-Bench}$  can be rejected, the benchmark model is not being encompassed and does explain variation of the target time series beyond the PLS model. Note, that both null hypotheses may be rejected implying that both models hold merit and may be combined. In contrast, one model is said to be encompassed by the competing model if its own null hypothesis can not be rejected, while that of the competing model is rejected. In other words, the encompassing model solely relies on the competing models while the other is disregarded.



## 4 Results

### 4.1 Unilateral Imports

This section reports the results for unilateral imports of ten representative countries, as well as world imports in total. Table 1 quantifies the forecast errors for the current month ( $h=0$ ) and one month ahead ( $h=1$ ) using the expanding window, pseudo out-of-sample approach. The columns labeled RMSFE display the absolute forecast error of the PLS model ("PLS, Abs."), while the other columns display the relative RMSFE with respect to the standard deviation (naive model) and the ARX benchmark model, respectively. The data can be interpreted as follows: For Brazilian seasonally and price adjusted imports in the current month ( $h=0$ ) the nowcast on the 30th day of the month predicts the month-over-month growth rate with an error of 6.3 percentage points. This amounts to approximately 67.9% of the standard deviation of Brazilian imports and to 65.5% of the ARX benchmark model's error. These relative improvements of the PLS model over other models is highly statistically significant as per the Diebold-Mariano test. As exemplified in this table and comprehensively reported in Tables 7 and 8, the PLS model's forecast errors lie below the standard deviation of all 76 import time series, except for Cyprus and North Macedonia. For the vast majority of countries, PLS also outperforms the benchmark model – frequently with high statistical significance. For twelve countries or regions marked in gray, the benchmark model generates lower forecast errors, although this difference is never significant. World imports, for instance, are well captured by the ARX benchmark model.

The encompassing tests in Table 2, as well as the comprehensive Tables 9 and 10 show the results of the encompassing test. The table can be interpreted as follows: The Wald test rejects the null hypothesis for the PLS model of German imports in  $h=0$ , namely that the PLS model is irrelevant for the combination model, with a p-value of 0.031. The PLS model is never encompassed by the ARX benchmark model for any of the countries and regions, although the null hypotheses for both the PLS and the ARX benchmark model cannot be rejected for Romanian imports one month ahead. In other words, the

PLS model always contributes unique predictive power to the combination model, save for the exception above.

Table 1: Forecast performance of PLS model for unilateral, real imports

	h=0			h=1		
	RMSFE PLS, Abs.	rel. RMSFE Std. Dev.	rel. RMSFE Benchmark	RMSFE PLS, Abs.	rel. RMSFE Std. Dev.	rel. RMSFE Benchmark
Brazil	0.063	0.679***	0.655***	0.065	0.716***	0.723***
China	0.036	0.690*	0.564**	0.035	0.671**	0.664**
EU	0.029	0.704	0.915	0.039	0.960	0.867
France	0.047	0.728*	0.705*	0.052	0.803	0.783
Germany	0.029	0.854	0.983	0.032	0.917***	0.838
Japan	0.034	0.825	0.774*	0.030	0.800	0.829
Russia	0.031	0.819	0.688**	0.031	0.804*	0.771*
Sub-S. Afr.	0.047	0.880	0.834	0.042	0.814***	0.748
UK	0.052	0.694**	0.684**	0.055	0.745**	0.751**
USA	0.021	0.688	0.835	0.022	0.739	0.708*
World	0.020	0.652	1.030	0.025	0.790*	0.709*

RMSFE refers to root mean squared forecast error. PLS, Abs. denotes the absolute RMSFE of the PLS model in decimal units. Rel. RMSFE Std. Dev and rel. RMSFE Benchmark denote the relative RMSFEs (13) of the PLS model divided by the standard deviation and the RMSFE of the ARX Benchmark model, respectively. A value below 1 indicates lower forecast errors of the PLS model. Relative RMSFE values above one marked in gray. Benchmark models are ARX-models using economic indicators applicable to the country. Significance levels calculated using Diebold-Mariano test. Sub-S. Afr. designates Sub-Saharan Africa. Cutoff day for PLS = 30.

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 2: Encompassing tests of PLS and benchmark models for unilateral, real imports

	h=0		h=1	
	Wald p-Value $H_0^{PLS} : \lambda_1 = 0$	Wald p-Value $H_0^{ARX-B.} : \lambda_2 = 0$	Wald p-Value $H_0^{PLS} : \lambda_1 = 0$	Wald p-Value $H_0^{ARX-B.} : \lambda_2 = 0$
Brazil	0.000***	0.319	0.000***	0.093*
China	0.000***	0.466	0.000***	0.685
EU27	0.000***	0.000***	0.071*	0.968
France	0.000***	0.479	0.000***	0.272
Germany	0.031**	0.027**	0.012**	0.803
Japan	0.000***	0.160	0.000***	0.042**
Russia	0.000***	0.887	0.000***	0.470
Sub-S. Afr.	0.000***	0.109	0.000***	0.351
UK	0.000***	0.144	0.000***	0.191
USA	0.000***	0.000***	0.000***	0.276
World	0.000***	0.000***	0.000***	0.800

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (15) and previous for definitions. Cells in gray show encompassed models: They do not explain variation of the combination model, while the competing model does. Cutoff day for PLS = 30.

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

## 4.2 Exports

Results for unilateral, real exports generally track previous ones for imports, although absolute forecast errors tend to be slightly higher. For the European Union, Germany and the UK in Table 3, for instance, absolute RMSFE values of the PLS model are around a percentage point higher than for respective imports. Except for the exports of Taiwan, the PLS models' forecast errors lie below the standard deviation in the current month,  $h = 0$ . The shipping model also frequently outperforms the ARX benchmark with high significance as seen in Tables 11 and 12. While the forecast errors of the benchmark model are lower than that of the PLS model for 16 countries, these differences are never significant. The relative performance of the PLS model tends to improve for the forecast one month ahead. The encompassing tests of Tables 4, as well as Tables 13 and 14 show that the PLS model for unilateral exports in the current month is encompassed only for two out of the 76 countries and regions: Latvia and Romania. Somewhat surprisingly, both null hypotheses of the PLS and the ARX benchmark model can not be rejected for the European Union's and the euro area's exports one month ahead. Note, however, that the p-values for extra EU exports are smaller than those for intra EU exports. It is conceivable that the intra European trade on truck, train and airplane one month ahead is poorly predicted by current container ship movements today.

Table 3: Forecast performance of PLS model for unilateral, real exports

	h=0			h=1		
	RMSFE PLS, Abs.	rel. RMSFE Std. Dev.	rel. RMSFE Benchmark	RMSFE PLS, Abs.	rel. RMSFE Std. Dev.	rel. RMSFE Benchmark
Brazil	0.055	0.823**	0.759**	0.056	0.855	0.851
China	0.053	0.791	0.782*	0.050	0.732**	0.731**
EU	0.039	0.775	1.013	0.050	0.998	0.907
France	0.055	0.807*	0.744*	0.066	0.974	0.961*
Germany	0.040	0.763	0.772	0.049	0.951	0.887
Japan	0.030	0.703**	0.870	0.033	0.770***	0.783**
Russia	0.047	0.741**	0.641***	0.053	0.822*	0.753*
Sub-S. Afr.	0.094	0.637	0.620	0.128	0.869	0.751
UK	0.059	0.856	0.805	0.052	0.785	0.744*
USA	0.040	0.746*	1.437	0.050	0.928	0.855**
World	0.022	0.683	1.219	0.031	0.947*	0.882

RMSFE refers to root mean squared forecast error. PLS, Abs. denotes the absolute RMSFE of the PLS model in decimal units. Rel. RMSFE Std. Dev and rel. RMSFE Benchmark denote the relative RMSFEs (13) of the PLS model divided by the standard deviation and the RMSFE of the ARX Benchmark model, respectively. A value below 1 indicates lower forecast errors of the PLS model. Relative RMSFE values above one marked in gray. Benchmark models are ARX-models using economic indicators applicable to the country. Significance levels calculated using Diebold-Mariano test. Sub-S. Afr. designates Sub-Saharan Africa. Cutoff day for PLS = 30.

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 4: Encompassing tests of PLS and benchmark models for unilateral, real exports

	h=0		h=1	
	Wald p-Value $H_0^{PLS} : \lambda_1 = 0$	Wald p-Value $H_0^{ARX-B.} : \lambda_2 = 0$	Wald p-Value $H_0^{PLS} : \lambda_1 = 0$	Wald p-Value $H_0^{ARX-B.} : \lambda_2 = 0$
Brazil	0.000***	0.204	0.000***	0.453
China	0.000***	0.751	0.000***	0.928
EU27	0.000***	0.000***	0.279	0.900
France	0.000***	0.769	0.073*	0.460
Germany	0.000***	0.774	0.052*	0.730
Japan	0.000***	0.012**	0.000***	0.418
Russia	0.000***	0.187	0.000***	0.479
Sub-S. Afr.	0.000***	0.730	0.002***	0.457
UK	0.000***	0.361	0.000***	0.511
USA	0.001***	0.000***	0.022**	0.625
World	0.000***	0.000***	0.059*	0.933

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (15) and previous for definitions. Cells in gray show encompassed models: They do not explain variation of the combination model, while the competing model does. Sub-S. Afr. designates Sub-Saharan Africa. Cutoff day for PLS = 30.

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

### 4.3 Bilateral

The flexible setup of the PLS algorithm also allows forecasting of specific bilateral trade flows. To exemplify this, PLS is trained on the bilateral imports of China, the European Union and the United States vis-à-vis their five largest trade partners. Table 5 reports the absolute and relative RMSFE values. Remarkably, the relative RMSFE against the standard deviation tends to be lower for the bilateral trade flows than for the unilateral imports. For instance, the RMSFE of the shipping model for the EU's imports from China, the United States, Russia and Switzerland ranges from 48 to 59% of the standard deviation. Forecast performance is also strong for China's imports. This may be the result of China's reliance on the container shipping network. In contrast, the EU's imports from the United Kingdom, as well as the US' imports from Mexico and Canada are less reliable in comparison to the benchmark. According to Table 6, however, the PLS model is only encompassed for the US imports from Canada.

Table 5: Forecast performance of PLS model for bilateral trade

	h=0			h=1		
	RMSFE PLS, Abs.	rel. RMSFE Std. Dev.	rel. RMSFE Benchmark	RMSFE PLS, Abs.	rel. RMSFE Std. Dev.	rel. RMSFE Benchmark
CN from AU	0.065	0.622***	0.586***	0.071	0.688***	0.676**
CN from DE	0.056	0.745***	0.720***	0.049	0.684***	0.650***
CN from JP	0.025	0.587***	0.373***	0.022	0.598***	0.382***
CN from KR	0.052	0.716***	0.526**	0.045	0.932**	0.855**
CN from US	0.053	0.675**	0.590***	0.048	0.750***	0.661***
EU from CN	0.035	0.553	0.341*	0.040	0.654	0.592*
EU from RU	0.066	0.591	0.563*	0.085	0.849*	0.832***
EU from CH	0.032	0.475**	0.703**	0.027	0.456**	0.462*
EU from UK	0.045	1.171**	0.410*	0.041	1.054***	0.979
EU from US	0.034	0.546*	0.560**	0.043	0.766*	0.709*
US from CA	0.058	0.680	0.917	0.071	0.837*	0.808**
US from CN	0.062	0.729	0.582*	0.066	0.797*	0.793**
US from DE	0.015	0.595*	0.447***	0.012	0.501**	0.389***
US from JP	0.031	0.692*	0.651*	0.034	0.638	0.589**
US from MX	0.083	0.879	1.219**	0.078	0.828	0.805

RMSFE refers to root mean squared forecast error. PLS, Abs. denotes the absolute RMSFE of the PLS model in decimal units. Rel. RMSFE Std. Dev and rel. RMSFE Benchmark denote the relative RMSFEs (13) of the PLS model divided by the standard deviation and the RMSFE of the ARX Benchmark model, respectively. A value below 1 indicates lower forecast errors of the PLS model. Relative RMSFE values above one marked in gray. Benchmark models are ARX-models using economic indicators applicable to the country. Significance levels calculated using Diebold-Mariano test. Cutoff day for PLS = 30.

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 6: Encompassing tests of PLS and benchmark models for bilateral trade

	h=0		h=1	
	Wald p-Value $H_0^{PLS} : \lambda_1 = 0$	Wald p-Value $H_0^{ARX-B} : \lambda_2 = 0$	Wald p-Value $H_0^{PLS} : \lambda_1 = 0$	p-Value $H_0^{ARX-B} : \lambda_2 = 0$
CN from AU	0.000***	0.447	0.001***	0.530
CN from DE	0.015**	0.852	0.000***	0.495
CN from JP	0.000***	0.781	0.000***	0.024**
CN from KR	0.000***	0.883	0.030**	0.578
CN from US	0.002***	0.375	0.011**	0.907
EU from CN	0.011**	0.368	0.000***	0.007***
EU from RU	0.000***	0.099*	0.001***	0.903
EU from CH	0.005***	0.274	0.000***	0.396
EU from UK	0.081*	0.082*	0.090*	0.191
EU from US	0.000***	0.728	0.003***	0.781
US from CA	0.244	0.096*	0.001***	0.056*
US from CN	0.005***	0.216	0.001***	0.139
US from DE	0.000***	0.261	0.000***	0.642
US from JP	0.001***	0.866	0.000***	0.231
US from MX	0.075*	0.000***	0.000***	0.501

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (15) and previous for definitions. Cells in gray show encompassed models: They do not explain variation of the combination model, while the competing model does. Cutoff day for PLS = 30.

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

#### 4.4 Performance During Crisis

In times of crisis, policy makers are in great need of reliable data to steer the economy. On the other hand, largely unexpected shocks such as the COVID-19 crisis make it more difficult to pinpoint the exact magnitude of disruptions. To study the properties of the shipping indicator in times of crisis, its performance is compared to autoregression (AR)<sup>2</sup> models for two time periods. The first period spans the five years before the Coronavirus crisis from 2015 to 2019. The second period focuses on the first half 2020, which saw the greatest disruptions in trade due to the pandemic. To illustrate the difference between the autoregression model and the PLS nowcasts  $\widehat{y}_{t+0}^{PLS}$ , the respective root mean square forecast errors are subtracted to reflect the excess forecast error of the AR model:

$$Excess\ AR\ Forecast\ Error = RMSFE_{h=0}^{AR} - RMSFE_{h=0}^{PLS} \quad (19)$$

<sup>2</sup>The setup is equivalent to the ARX specification (11) without the leading indicators

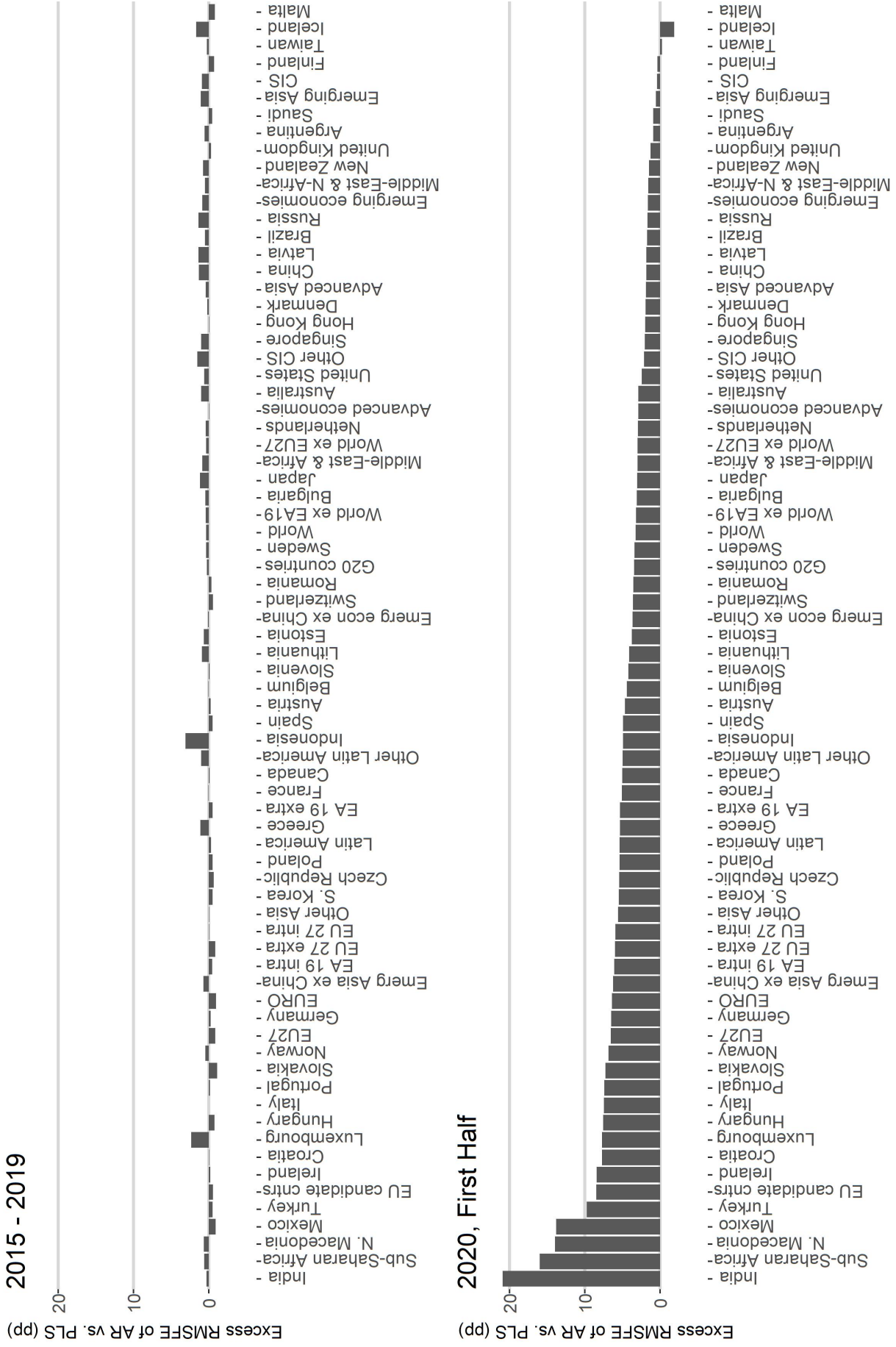
A positive value, therefore, signifies that the container shipping forecast outperforms the AR model. Figure 3 summarizes the excess forecast error of AR models as compared to the PLS system. In the five years preceeding the COVID-19 crisis, the AR model produces forecast errors that exceed those of the PLS system by 0 to 3 percentage points for the majority of countries. For several countries such as Mexico, Turkey and the euro area the AR model performs slightly better. In the first half of 2020, the excess forecast error of AR models jumps for essentially all countries and stands above 10 percentage points for major economies such as India and Mexico. Furthermore, the excess error of AR models ranges between 5 to 10 percentage points for almost half the countries in the sample. This shows that the PLS model tends to be more reliable than models based on autoregression terms in times of economic turmoil. This holds both for developed and developing countries.

## 5 Discussion

A unique feature of the approach outlined above is that the forecast can be calculated at any point of the month. If not stated otherwise, above results rely on ship positions and port calls up to the 30th day of the month. The first robustness check in Appendix A.2 tests the PLS indicator for unilateral imports and exports using the 30 days period leading up to the 15th of the month. This allows publication of the PLS forecasts well before the end of the month. As Tables 15 and 17 show, the absolute forecast errors for the sample economies rise slightly. For instance, it rises by 0.2 percentage points for German imports and 1.4 percentage points for French imports. Nevertheless, the relative RMSFE against the ARX benchmark model never increases over 1 with statistical significance. The PLS estimates are also never encompassed as seen in Tables 16 and 18.

The second robustness check estimates the growth rate of nominal unilateral trade flows instead of price adjusted ones in the main specifications. Conceivably, the movement of container ships only reflect the growth of trade volume, but not the growth of trade value. This is not generally the case. The above exercises are repeated for the forecasts

Figure 3: Excess Forecast Error of AR model vs. new shipping indicator (PLS) before and during COVID-19 crisis<sup>a</sup>



<sup>a</sup>Two countries omitted for graphical reasons: South Africa (2015-2019: -2.0; 2020, First Half 2020: 28.6) and Cyprus (2015-2019: 10.2, First Half 2020: -7.7)



of nominal, unilateral trade flows obtained from the World Trade Organisation. Again, the absolute forecast errors tend to rise slightly as reported in Tables 19 and 21. This trend is less clear for the relative RMSFEs. For example, the relative RMSFEs against the standard deviation and the ARX benchmark model fall for Germany, the EU and Russia. This indicates that the higher volatility of nominal trade flows causes the higher absolute RMSFEs.

Potentially the largest threat to the predictive power remains overfitting of the PLS model, i.e. the possibility that the PLS model heavily relies on explanatory variables that correlate with the dependent trade time series by chance. Indeed, reducing the number of explanatory variables by merging small ports in close proximity and applying the VIP method to select 30 final explanatory variables, significantly boosts the forecast quality in the out-of-sample exercises. It may nevertheless be possible that explanatory time series that coincidentally show correlation "survive" the above filtering techniques. This would impair the identification of appropriate latent variables and result in larger forecast errors in above out-of-sample studies. Nevertheless, the PLS model appears to perform well. Whatever overfitting remains, it does not render the PLS model inadequate. In contrast, exploratory analyses showed that the PLS model performs better in this setting than the related Principle Component Regression (PCR). If overfitting was an issue, one would expect the opposite. This is because PLS exploits covariance between dependent and independent time series in calculation of the factors. Hence, PLS would arguably be more prone to overfitting than PCR. Lastly, visual inspection of selected variables does not indicate heavy overfitting. The thirty variables selected for forecasting US imports are reported in Appendix A.4. One might expect the algorithm to select both US American ports and major world ports to reflect US American consumer demand, as well world economic activity. Indeed, eight of the thirty variables are located directly in the US. Other variables track ship movements of the coast of Mexico or developments of major economies such as Germany, India and Japan.

Other limitations of this forecasting exercise may include structural breaks. For instance, the Corona crisis has changed, at least for a period in 2020, the importance of

container shipping for trade in comparison to other modes of transport such as trade by aircraft. This may lead the PLS algorithm to over- or underestimate the change of real trade flows. One could hypothetically investigate this potential structural break by comparing the forecast quality of a model trained without the Corona crisis with a model trained including the crisis. Unfortunately, this is impossible to perform with the currently available 82 monthly observations between 2015 and mid-2021. In light of possible overfitting, PLS must be trained using at least 40 observations. Establishing significance levels with the Diebold-Mariano test requires an additional 30 to 40 observations. These requirements in effect lock the time of the Corona crisis in the testing phase. Overall, the effect of potential breaks does not appear to be very detrimental. After all, the section on results establishes the high predictive power while the out-of-sample period always includes the Corona crisis. Practitioners may nevertheless combine forecasts from the container shipping network with leading indicators derived from other sources. One natural extension of this work, therefore, includes integration of other modes of transportation such as flight recordings or toll information on truck traffic.

## 6 Conclusion

This paper introduces a sparse partial least squares algorithm that estimates month-over-month growth of price adjusted trade flows from high-frequency container ship movements. The method can calculate forecasts for unilateral imports and exports, as well as bilateral trade flows of any country and economic bloc in the world. Although the time spans to train the algorithm and to test out-of-sample predictions are relatively short, results indicate lower forecast errors of the new model as compared to benchmark time series models. Encompassing tests furthermore show that the forecasts from the partial least squares model contribute meaningful predictive power to combination models. This also holds true for nominal trade flows, as well as nowcasts calculated during an ongoing month. While overfitting and structural breaks may slightly deteriorate the forecasts, the out-of-sample exercises continue to show significant predictive power. As more monthly

observations in the shipping data become available, these concerns can be addressed more rigorously. Extensions of this work may include an application of the PLS model to other sources of transportation data such as flight positions or truck tolls. Additionally, the currently available methods and data can be trained on other dependent time series beyond national headline trade figures. The flexibility of factor models also would allow the prediction of trade and economic activity of subnational regions or specific industries.

## References

- Arslanalp, S., Marini, M., and Tumbarello, P. (2019). Big data on vessel traffic: Nowcasting trade flows in real time.
- Bai, J. and Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70(1):191–221.
- Bai, J. and Ng, S. (2004). A panic attack on unit roots and cointegration. *Econometrica*, 72(4):1127–1177.
- Boivin, J. and Ng, S. (2006). Are more data always better for factor analysis? *Journal of Econometrics*, 132(1):169–194.
- Brancaccio, G., Kalouptsi, M., and Papageorgiou, T. (2017). Geography, search frictions and endogenous trade costs. Technical report, National Bureau of Economic Research.
- Breitung, J. and Knüppel, M. (2021). How far can we forecast? statistical tests of the predictive content. *Journal of Applied Econometrics*, 36(4):369–392.
- Bureau, U. C. (2017). X-13arima-seats reference manual.
- Camacho, M. and Perez-Quiros, G. (2010). Introducing the euro-sting: Short-term indicator of euro area growth. *Journal of Applied Econometrics*, 25(4):663–694.
- Celasun, O., Hansen, N.-J., Mineshima, A., Spector, M., and Zhou, J. (2022). Supply bottlenecks: Where, why, how much, and what next? Technical report.
- Cerdeiro, D. A., Komaromi, A., Liu, Y., Saeed, M., et al. (2020). World seaborne trade in real time: A proof of concept for building ais-based nowcasts from scratch. Technical report.
- Chun, H. and Keleş, S. (2010). Sparse partial least squares regression for simultaneous dimension reduction and variable selection. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72(1):3–25.
- Cocchi, M., Biancolillo, A., and Marini, F. (2018). Chapter ten—chemometric methods for classification and feature selection. *Comprehensive Analytical Chemistry; Jaumot, J., Bedia, C., Tauler, R., Eds*, 82:265–299.
- CPB (2021). Cpb netherlands bureau for economic policy analysis. world trade monitor. Database accessed in February 2021, <https://www-genesis.destatis.de/genesis/online>.
- Davidson, R., MacKinnon, J. G., et al. (1993). *Estimation and inference in econometrics*, volume 63. Oxford New York.
- DESTATIS (2021). Statistisches bundesamt. various leading indicators. Auftragseingänge, Auftragsbestand and LKW-Maut-Fahrleistungsindex, <https://www-genesis.destatis.de/genesis/online>.
- Diebold, F. X. and Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & economic statistics*, 20(1):134–144.

- Döhrn, R. and Maatsch, S. (2012). Der rwi/isl-containerumschlag-index. *Wirtschaftsdienst*, 92(5):352–354.
- Eickmeier, S. and Ng, T. (2011). Forecasting national activity using lots of international predictors: An application to new zealand. *International Journal of Forecasting*, 27(2):496–511.
- Esposito Vinzi, V., Chin, W. W., Henseler, J., Wang, H., et al. (2010). *Handbook of partial least squares*, volume 201. Springer.
- European Commission (2020). Business and consumer surveys. Main Indicators, [https://ec.europa.eu/info/business-economy-euro/indicators-statistics/economic-databases/business-and-consumer-surveys/download-business-and-consumer-survey-data/time-series\\_en](https://ec.europa.eu/info/business-economy-euro/indicators-statistics/economic-databases/business-and-consumer-surveys/download-business-and-consumer-survey-data/time-series_en).
- Federal Reserve Bank of St. Louis (2020). Board of governors of the federal reserve system (us), industrial production. retrieved from FRED in February 2021, <https://fred.stlouisfed.org/series/INDPRO>.
- Fuentes, J., Poncela, P., and Rodríguez, J. (2015). Sparse partial least squares in time series for macroeconomic forecasting. *Journal of Applied Econometrics*, 30(4):576–595.
- Galindo-Prieto, B., Eriksson, L., and Trygg, J. (2014). Variable influence on projection (vip) for orthogonal projections to latent structures (opls). *Journal of Chemometrics*, 28(8):623–632.
- Garthwaite, P. H. (1994). An interpretation of partial least squares. *Journal of the American Statistical Association*, 89(425):122–127.
- Grimme, C., Lehmann, R., and Noeller, M. (2018). Forecasting imports with information from abroad.
- Grimme, C. and Wohlrabe, K. (2014). Die ifo exporterwartungen—ein neuer indikator zur lage der exportindustrie in deutschland. *ifo Schnelldienst*, 67(23):64–65.
- Haenlein, M. and Kaplan, A. M. (2004). A beginner’s guide to partial least squares analysis. *Understanding statistics*, 3(4):283–297.
- Hartigan, J. A. and Wong, M. A. (1979). Algorithm as 136: A k-means clustering algorithm. *Journal of the royal statistical society. series c (applied statistics)*, 28(1):100–108.
- Heiland, I., Moxnes, A., Ulltveit-Moe, K. H., and Zi, Y. (2019). Trade from space: Shipping networks and the global implications of local shocks. *Working Paper*.
- Ifo institute (2021). Various leading indicators. Ifo Zeitreihen - Exporterwartungen, Import- and Exportklima, <https://www.ifo.de/umfragen/zeitreihen>.
- IMF (2021). International monetary fund. dots. data retrieved from the Direction of Trade Statistics in February 2021, <https://data.imf.org/regular.aspx?key=61726508>.
- Jurado, K., Ludvigson, S. C., and Ng, S. (2015). Measuring uncertainty. *American Economic Review*, 105(3):1177–1216.

- Lohmöller, J.-B. (2013). *Latent variable path modeling with partial least squares*. Springer Science & Business Media.
- Rönkkö, M. (2014). The effects of chance correlations on partial least squares path modeling. *Organizational Research Methods*, 17(2):164–181.
- RWI/ISL (2020). Containerumschlag-index. retrieved from RWI Essen in February 2021, <https://www.rwi-essen.de/containerindex>.
- Stock, J. H. and Watson, M. W. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American statistical association*, 97(460):1167–1179.
- Tran, T. N., Afanador, N. L., Buydens, L. M., and Blanchet, L. (2014). Interpretation of variable importance in partial least squares with significance multivariate correlation (smc). *Chemometrics and Intelligent Laboratory Systems*, 138:153–160.
- UNCTAD (2021). Review of maritime transport 2021.
- Wold, H. (1975). Path models with latent variables: The nipals approach. In *Quantitative sociology*, pages 307–357. Elsevier.
- Wold, S., Ruhe, A., Wold, H., and Dunn, Iii, W. (1984). The collinearity problem in linear regression. the partial least squares (pls) approach to generalized inverses. *SIAM Journal on Scientific and Statistical Computing*, 5(3):735–743.
- Wold, S., Sjöström, M., and Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemometrics and intelligent laboratory systems*, 58(2):109–130.
- Wong, W. F. and Ziv, O. (2020). Entrepã´ t: Hubs, scale, and trade costs.
- WTO (2020). World trade organization. wto data. data retrieved from data portal in February 2021, <https://data.wto.org/>.



# A Appendix

## A.1 Full Country Results

Table 7: Forecast performance of PLS model for unilateral, real imports (I/II)

	h=0			h=1		
	RMSFE PLS, Abs.	rel. RMSFE Std. Dev.	rel. RMSFE Benchmark	RMSFE PLS, Abs.	rel. RMSFE Std. Dev.	rel. RMSFE Benchmark
Argentina	0.063	0.590**	0.609**	0.065	0.652***	0.588***
Australia	0.029	0.583***	0.466***	0.045	0.892	0.928
Austria	0.030	0.701*	0.772**	0.035	0.820	0.765
Belgium	0.031	0.677*	0.744**	0.035	0.769*	0.712*
Brazil	0.063	0.679***	0.655***	0.065	0.716***	0.723***
Bulgaria	0.038	0.685*	0.820*	0.046	0.854	0.870
Canada	0.042	0.692*	1.182	0.049	0.788	0.726
China	0.036	0.690*	0.564**	0.035	0.671**	0.664**
Croatia	0.052	0.720**	0.848	0.051	0.701	0.650*
Cyprus	0.177	1.017	1.065	0.158	0.895	0.916
Czech Rep.	0.034	0.725*	0.740**	0.035	0.756	0.715
Denmark	0.026	0.732	0.662*	0.025	0.718	0.686
Estonia	0.034	0.694	0.669*	0.049	0.970	0.885
Finland	0.028	0.634*	0.662**	0.033	0.768*	0.714**
France	0.047	0.728*	0.705*	0.052	0.803	0.783
Germany	0.029	0.854	0.983	0.032	0.917***	0.838
Greece	0.057	0.861	0.752	0.055	0.806*	0.808
Hong Kong	0.023	0.711***	0.450**	0.021	0.655**	0.591***
Hungary	0.034	0.673	0.687**	0.041	0.799	0.784
Iceland	0.072	0.680*	0.803	0.069	0.653*	0.628**
India	0.072	0.589*	0.588*	0.102	0.844	0.794
Indonesia	0.099	0.732**	0.806	0.100	0.567**	0.460
Ireland	0.075	0.680**	0.752	0.083	0.753**	0.722***
Italy	0.043	0.800*	0.765	0.047	0.881	0.856
Japan	0.034	0.825	0.774*	0.030	0.800	0.829
S. Korea	0.033	0.776***	0.730***	0.026	0.612***	0.596***
Latvia	0.034	0.654*	0.854**	0.041	0.790**	0.786**
Lithuania	0.030	0.658	1.042	0.038	0.839	0.776
Luxemb.	0.063	0.760*	0.683**	0.060	0.742*	0.747*
Malta	0.158	0.649***	0.720**	0.169	0.727*	0.698
Mexico	0.052	0.746*	0.864	0.049	0.691	0.569
Netherl.	0.023	0.702**	0.715**	0.026	0.813***	0.759**
New Zeal.	0.036	0.621***	0.807*	0.037	0.637***	0.636***
N. Maced.	0.012	1.186	0.467***	0.021	1.954**	0.968
Norway	0.049	0.760***	0.665	0.053	0.792**	0.755***
Poland	0.037	0.713	1.082	0.044	0.858	0.864
Portugal	0.055	0.710*	0.777**	0.058	0.771	0.757
Romania	0.060	0.874	1.027	0.069	0.999	0.972
Russia	0.031	0.819	0.688**	0.031	0.804*	0.771*

RMSFE refers to root mean squared forecast error. PLS, Abs. denotes the absolute RMSFE of the PLS model in decimal units. Rel. RMSFE Std. Dev and rel. RMSFE Benchmark denote the relative RMSFEs (13) of the PLS model divided by the standard deviation and the RMSFE of the ARX Benchmark model, respectively. A value below 1 indicates lower forecast errors of the PLS model. Relative RMSFE values above one marked in gray. Benchmark models are ARX-models using economic indicators applicable to the country. Significance levels calculated using Diebold-Mariano test. N. Maced. designates North Macedonia. Cutoff day for PLS = 30.

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01



Table 8: Forecast performance of PLS model for unilateral, real imports (II/II)

	h=0			h=1		
	RMSFE PLS, Abs.	rel. RMSFE Std. Dev.	rel. RMSFE Benchmark	RMSFE PLS, Abs.	rel. RMSFE Std. Dev.	rel. RMSFE Benchmark
Adv. Asia	0.024	0.816**	0.617*	0.020	0.667***	0.658***
Adv. econ.	0.020	0.667	1.215	0.025	0.833	0.764
CIS	0.028	0.805	0.708*	0.027	0.769**	0.717**
EA 19 extra	0.023	0.688**	0.722*	0.028	0.839**	0.738*
EA 19 intra	0.045	0.874	1.239	0.048	0.933	0.886
EU 27 extra	0.023	0.720*	0.731	0.026	0.817**	0.725**
EU 27 intra	0.044	0.863	1.236	0.048	0.941	0.894
EU candid.	0.041	0.611	0.652*	0.053	0.786	0.751
EU27	0.029	0.704	0.915	0.039	0.960	0.867
EURO	0.030	0.732*	0.949	0.039	0.943	0.844
EM A. exCN	0.037	0.672	0.916	0.049	0.857	0.800*
EM exCN	0.031	0.731	0.981	0.034	0.787*	0.700
EM Asia	0.030	0.729**	0.785*	0.031	0.761	0.766
EM	0.027	0.750	0.889	0.028	0.770**	0.711**
G20 countr.	0.020	0.651	1.159	0.027	0.860	0.784*
Latin Am.	0.037	0.701*	0.767	0.040	0.774	0.712
M.-E./ Afr.	0.037	0.759	0.907	0.034	0.786**	0.638
M.-E./ N.Afr.	0.043	0.779	0.930	0.035	0.730**	0.610*
Other Asia	0.031	0.610***	0.693**	0.033	0.654**	0.584**
Other CIS	0.036	0.739	0.768*	0.038	0.810**	0.792**
O. Latin Am.	0.038	0.766*	0.795**	0.041	0.848	0.834**
Saudi Arab.	0.069	0.663**	0.788***	0.065	0.645**	0.502*
Singapore	0.034	0.855	0.838	0.028	0.713***	0.662***
Slovak Rep.	0.055	0.796	0.796*	0.055	0.797	0.806
Slovenia	0.042	0.580*	0.604**	0.052	0.704	0.652
South Afr.	0.055	0.706**	0.565*	0.060	0.819	0.796*
Spain	0.045	0.721	0.752*	0.056	0.906	0.888*
Sub-S. Afr.	0.047	0.880	0.834	0.042	0.814***	0.748
Sweden	0.024	0.656*	0.683**	0.032	0.886***	0.804**
Switzerland	0.029	0.581	0.842*	0.030	0.665	0.638
Taiwan	0.031	0.774**	0.547	0.034	0.837**	0.816**
Turkey	0.041	0.611	0.635*	0.053	0.786	0.738
UK	0.052	0.694**	0.684**	0.055	0.745**	0.751**
USA	0.021	0.688	0.835	0.022	0.739	0.708*
World exEA	0.021	0.753	1.130	0.022	0.793*	0.719**
World exEU	0.020	0.742	1.030	0.022	0.792	0.718*
World	0.020	0.652	1.030	0.025	0.790*	0.709*

RMSFE refers to root mean squared forecast error. PLS, Abs. denotes the absolute RMSFE of the PLS model in decimal units. Rel. RMSFE Std. Dev and rel. RMSFE Benchmark denote the relative RMSFEs (13) of the PLS model divided by the standard deviation and the RMSFE of the ARX Benchmark model, respectively. A value below 1 indicates lower forecast errors of the PLS model. Relative RMSFE values above one marked in gray. Benchmark models are ARX-models using economic indicators applicable to the country. Significance levels calculated using Diebold-Mariano test. EA refers to euro area, EM to emerging economies, exCN to ex China, exEA to ex euro area, exEU to ex EU27, M.-E./ (N.) Afr. to Middle-East and (North) Africa, O. Latin Am. to Other Latin America and Sub-S. Afr. designates Sub-Saharan Africa. Cutoff day for PLS = 30.

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 9: Encompassing tests of PLS and benchmark models for unilateral, real imports (I/II)

	h=0		h=1	
	Wald p-Value $H_0^{PLS} : \lambda_1 = 0$	Wald p-Value $H_0^{ARX-B.} : \lambda_2 = 0$	Wald p-Value $H_0^{PLS} : \lambda_1 = 0$	p-Value $H_0^{ARX-B.} : \lambda_2 = 0$
Argentina	0.000***	0.052*	0.000***	0.238
Australia	0.000***	0.880	0.010**	0.289
Austria	0.000***	0.004***	0.000***	0.979
Belgium	0.000***	0.083*	0.000***	0.786
Brazil	0.000***	0.319	0.000***	0.093*
Bulgaria	0.000***	0.006***	0.000***	0.091*
Canada	0.000***	0.000***	0.000***	0.731
China	0.000***	0.466	0.000***	0.685
Croatia	0.000***	0.014**	0.000***	0.218
Cyprus	0.024**	0.081*	0.001***	0.406
Czech Rep.	0.000***	0.445	0.000***	0.948
Denmark	0.000***	0.455	0.000***	0.421
Estonia	0.000***	0.631	0.004***	0.100
Finland	0.000***	0.586	0.000***	0.351
France	0.000***	0.479	0.000***	0.272
Germany	0.031**	0.027**	0.012**	0.803
Greece	0.000***	0.521	0.000***	0.838
Hong Kong	0.000***	0.871	0.000***	0.666
Hungary	0.000***	0.244	0.000***	0.066*
Iceland	0.000***	0.016**	0.000***	0.437
India	0.000***	0.149	0.001***	0.763
Indonesia	0.000***	0.221	0.000***	0.390
Ireland	0.000***	0.350	0.000***	0.903
Italy	0.000***	0.048**	0.003***	0.951
Japan	0.000***	0.160	0.000***	0.042**
South Korea	0.000***	0.574	0.000***	0.310
Latvia	0.000***	0.157	0.000***	0.517
Lithuania	0.004***	0.000***	0.000***	0.791
Luxembourg	0.000***	0.786	0.000***	0.704
Malta	0.000***	0.360	0.000***	0.721
Mexico	0.000***	0.004***	0.000***	0.616
Netherlands	0.000***	0.376	0.000***	0.680
New Zealand	0.000***	0.003***	0.000***	0.138
N. Macedonia	0.015**	0.043**	0.038**	0.066*
Norway	0.000***	0.866	0.000***	0.548
Poland	0.000***	0.000***	0.000***	0.268
Portugal	0.000***	0.122	0.000***	0.424
Romania	0.003***	0.000***	0.235	0.798
Russia	0.000***	0.887	0.000***	0.470

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (15) and previous for definitions. Cells in gray show encompassed models: They do not explain variation of the combination model, while the competing model does. Cutoff day for PLS = 15.

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 10: Encompassing tests of PLS and benchmark models for unilateral, real imports (II/II)

	h=0		h=1	
	Wald p-Value $H_0^{PLS} : \lambda_1 = 0$	Wald p-Value $H_0^{ARX-B.} : \lambda_2 = 0$	Wald p-Value $H_0^{PLS} : \lambda_1 = 0$	p-Value $H_0^{ARX-B.} : \lambda_2 = 0$
Advanced Asia	0.000***	0.799	0.000***	0.698
Advanced econ.	0.000***	0.000***	0.000***	0.805
CIS	0.000***	0.334	0.000***	0.812
EA 19 extra	0.000***	0.009***	0.000***	0.285
EA 19 intra	0.000***	0.000***	0.016**	0.402
EU 27 extra	0.000***	0.011**	0.000***	0.461
EU 27 intra	0.000***	0.000***	0.019**	0.360
EU candidates	0.000***	0.089*	0.000***	0.857
EU27	0.000***	0.000***	0.071*	0.968
EURO	0.000***	0.000***	0.041**	0.940
Em. Asia ex CN	0.000***	0.000***	0.000***	0.656
Em. econ ex CN	0.000***	0.000***	0.000***	0.887
EM Asia	0.000***	0.090*	0.000***	0.709
EM econ.	0.000***	0.008***	0.000***	0.874
G20 countries	0.000***	0.000***	0.000***	0.365
Latin America	0.000***	0.005***	0.000***	0.156
Mid.-E./ Afr.	0.001***	0.097*	0.000***	0.197
Mid.-E./ N-Afr.	0.002***	0.071*	0.000***	0.294
Other Asia	0.000***	0.166	0.000***	0.504
Other CIS	0.000***	0.065*	0.000***	0.966
Other Lat. Am.	0.000***	0.001***	0.000***	0.871
Saudi Arabia	0.000***	0.115	0.000***	0.294
Singapore	0.001***	0.004***	0.000***	0.525
Slovak Republic	0.000***	0.822	0.000***	0.071*
Slovenia	0.000***	0.200	0.000***	0.505
South Africa	0.000***	0.902	0.000***	0.220
Spain	0.000***	0.024**	0.009***	0.865
Sub-Sah. Afr.	0.000***	0.109	0.000***	0.351
Sweden	0.000***	0.187	0.001***	0.023**
Switzerland	0.000***	0.001***	0.000***	0.046**
Taiwan	0.000***	0.323	0.000***	0.936
Turkey	0.000***	0.065*	0.000***	0.770
UK	0.000***	0.144	0.000***	0.191
USA	0.000***	0.000***	0.000***	0.276
World ex EA19	0.005***	0.000***	0.000***	0.654
World ex EU27	0.001***	0.000***	0.000***	0.980
World	0.000***	0.000***	0.000***	0.393

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (15) and previous for definitions. Cells in gray show encompassed models: They do not explain variation of the combination model, while the competing model does. EA refers to euro area, Em. econ to emerging economies, ex CN to ex China, ex EA19 to ex euro area, Middle.-E./ (N.) Afr. to Middle-East and (North) Africa, and Sub-Sah. Afr. designates Sub-Saharan Africa. Cutoff day for PLS = 30.

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 11: Forecast performance of PLS model for unilateral, real exports (I/II)

	h=0			h=1		
	RMSFE PLS, Abs.	rel. RMSFE Std. Dev.	rel. RMSFE Benchmark	RMSFE PLS, Abs.	rel. RMSFE Std. Dev.	rel. RMSFE Benchmark
Argentina	0.074	0.765	0.742*	0.074	0.771	0.770
Australia	0.033	0.727***	0.682**	0.033	0.709**	0.662**
Austria	0.029	0.708	0.800*	0.038	0.893	0.822
Belgium	0.035	0.707	0.849	0.043	0.869*	0.795**
Brazil	0.055	0.823**	0.759**	0.056	0.855	0.851
Bulgaria	0.038	0.752***	0.728*	0.031	0.630***	0.549***
Canada	0.046	0.718	1.428	0.061	0.967	0.960
China	0.053	0.791	0.782*	0.050	0.732**	0.731**
Croatia	0.053	0.676**	0.769*	0.051	0.633**	0.499***
Cyprus	0.299	0.653	0.710	0.326	0.709	0.701
Czech Rep.	0.056	0.853	0.722*	0.052	0.800	0.769
Denmark	0.026	0.744***	0.778**	0.027	0.795**	0.780**
Estonia	0.034	0.671**	0.735**	0.034	0.689***	0.620**
Finland	0.060	0.806**	0.993	0.051	0.678*	0.648**
France	0.055	0.807*	0.744*	0.066	0.974	0.961*
Germany	0.040	0.763	0.772	0.049	0.951	0.887
Greece	0.046	0.777*	0.664**	0.049	0.848**	0.854*
Hong Kong	0.038	0.698**	0.719**	0.027	0.489**	0.497**
Hungary	0.053	0.747	0.722*	0.068	0.960	0.966
Iceland	0.082	0.732*	0.829	0.080	0.704**	0.683**
India	0.105	0.643	0.890	0.160	0.988	0.913
Indonesia	0.050	0.618***	0.761*	0.054	0.600**	0.472**
Ireland	0.060	0.568**	0.739	0.066	0.616*	0.609*
Italy	0.070	0.842	0.710	0.065	0.779	0.792
Japan	0.030	0.703**	0.870	0.033	0.770***	0.783**
S. Korea	0.039	0.726	0.897	0.039	0.740	0.726
Latvia	0.030	0.909	1.041	0.030	0.919	0.888*
Lithuania	0.030	0.624	0.765*	0.035	0.735	0.652
Luxemb.	0.059	0.706*	0.716*	0.072	0.840	0.823
Malta	0.128	0.827	1.045	0.127	0.840	0.820
Mexico	0.091	0.717	0.860	0.107	0.844	0.630
Netherl.	0.024	0.728	0.884	0.030	0.903*	0.832**
New Zeal.	0.028	0.645***	0.575*	0.034	0.796**	0.667**
N. Maced.	0.147	0.785	0.854	0.148	0.793	0.758
Norway	0.061	0.773*	0.731***	0.066	0.837**	0.807***
Poland	0.047	0.829	0.933	0.051	0.913	0.954
Portugal	0.056	0.726*	0.726**	0.074	0.953	0.964
Romania	0.090	0.970	1.103	0.088	0.947	1.000
Russia	0.047	0.741**	0.641***	0.053	0.822*	0.753*

RMSFE refers to root mean squared forecast error. PLS, Abs. denotes the absolute RMSFE of the PLS model in decimal units. Rel. RMSFE Std. Dev and rel. RMSFE Benchmark denote the relative RMSFEs (13) of the PLS model divided by the standard deviation and the RMSFE of the ARX Benchmark model, respectively. A value below 1 indicates lower forecast errors of the PLS model. Relative RMSFE values above one marked in gray. Benchmark models are ARX-models using economic indicators applicable to the country. Significance levels calculated using Diebold-Mariano test. N. Maced. designates North Macedonia. Cutoff day for PLS = 30.

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 12: Forecast performance of PLS model for unilateral, real exports (II/II)

	h=0			h=1		
	RMSFE PLS, Abs.	rel. RMSFE Std. Dev.	rel. RMSFE Benchmark	RMSFE PLS, Abs.	rel. RMSFE Std. Dev.	rel. RMSFE Benchmark
Adv. Asia	0.022	0.765*	0.770**	0.022	0.715	0.716*
Adv. econ.	0.031	0.787	1.626	0.040	1.027	0.997
CIS	0.042	0.788**	0.767*	0.044	0.839*	0.796*
EA 19 extra	0.037	0.749	1.139	0.047	0.933	0.879
EA 19 intra	0.042	0.801	1.119	0.054	1.043	0.996
EU 27 extra	0.035	0.731	1.127	0.045	0.945	0.888
EU 27 intra	0.045	0.859	1.163	0.054	1.034	0.983
EU candid.	0.064	0.721	0.876	0.071	0.796	0.736
EU27	0.039	0.775	1.013	0.050	0.998	0.907
EURO	0.040	0.778	0.990	0.050	0.989	0.867
EM A. exCN	0.031	0.591**	1.032	0.047	0.886**	0.848**
EM exCN	0.026	0.676	1.099	0.036	0.923	0.827
EM Asia	0.038	0.858	0.787**	0.032	0.715**	0.714**
EM	0.025	0.798*	0.755**	0.024	0.760**	0.684***
G20 countr.	0.027	0.735	1.329	0.034	0.927	0.895
Latin Am.	0.037	0.663	1.008	0.040	0.723	0.563
M.-E./ Afr.	0.027	0.762	0.617	0.030	0.845	0.808
M.-E./ N.Afr.	0.028	0.822	0.651	0.031	0.912	0.931*
Other Asia	0.030	0.649**	0.922	0.034	0.732***	0.731***
Other CIS	0.043	0.769***	0.572**	0.040	0.715**	0.734**
O. Latin Am.	0.028	0.681**	0.683**	0.032	0.771**	0.756**
Saudi Arab.	0.092	0.911	0.969	0.092	0.905	0.907
Singapore	0.031	0.777**	0.805**	0.033	0.809***	0.788***
Slovak Rep.	0.076	0.783	0.671*	0.079	0.815	0.798
Slovenia	0.044	0.810*	0.871	0.052	0.951	0.903
South Afr.	0.129	0.663	0.624	0.174	0.893	0.760
Spain	0.057	0.864	0.891	0.065	0.980	1.025
Sub-S. Afr.	0.094	0.637	0.620	0.128	0.869	0.751
Sweden	0.031	0.763**	0.884	0.035	0.856*	0.772***
Switzerland	0.040	0.753***	0.994	0.037	0.694***	0.670***
Taiwan	0.021	1.123	0.652	0.020	1.078	1.067
Turkey	0.059	0.677	0.564	0.070	0.793	0.753
UK	0.059	0.856	0.805	0.052	0.785	0.744*
USA	0.040	0.746*	1.437	0.050	0.928	0.855**
World exEA	0.021	0.696	0.995	0.025	0.840**	0.767**
World exEU	0.022	0.730	0.975	0.024	0.796***	0.727***
World	0.022	0.683	1.219	0.031	0.947*	0.882

RMSFE refers to root mean squared forecast error. PLS, Abs. denotes the absolute RMSFE of the PLS model in decimal units. Rel. RMSFE Std. Dev and rel. RMSFE Benchmark denote the relative RMSFEs (13) of the PLS model divided by the standard deviation and the RMSFE of the ARX Benchmark model, respectively. A value below 1 indicates lower forecast errors of the PLS model. Relative RMSFE values above one marked in gray. Benchmark models are ARX-models using economic indicators applicable to the country. Significance levels calculated using Diebold-Mariano test. EA refers to euro area, EM to emerging economies, exCN to ex China, exEA to ex euro area, exEU to ex EU27, M.-E./ (N.) Afr. to Middle-East and (North) Africa, O. Latin Am. to Other Latin America and Sub-S. Afr. designates Sub-Saharan Africa. Cutoff day for PLS = 30.

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 13: Encompassing tests of PLS and benchmark models for unilateral, real exports (I/II)

	h=0		h=1	
	Wald p-Value $H_0^{PLS} : \lambda_1 = 0$	Wald p-Value $H_0^{ARX-B.} : \lambda_2 = 0$	Wald p-Value $H_0^{PLS} : \lambda_1 = 0$	p-Value $H_0^{ARX-B.} : \lambda_2 = 0$
Argentina	0.000***	0.463	0.000***	0.917
Australia	0.000***	0.630	0.000***	0.554
Austria	0.000***	0.001***	0.006***	0.731
Belgium	0.000***	0.001***	0.005***	0.180
Brazil	0.000***	0.204	0.000***	0.453
Bulgaria	0.000***	0.101	0.000***	0.617
Canada	0.008***	0.000***	0.026**	0.111
China	0.000***	0.751	0.000***	0.928
Croatia	0.000***	0.232	0.000***	0.040**
Cyprus	0.000***	0.458	0.000***	0.047**
Czech Rep.	0.001***	0.020**	0.000***	0.674
Denmark	0.000***	0.888	0.000***	0.351
Estonia	0.000***	0.219	0.000***	0.128
Finland	0.001***	0.001***	0.000***	0.828
France	0.000***	0.769	0.073*	0.460
Germany	0.000***	0.774	0.052*	0.730
Greece	0.000***	0.042**	0.000***	0.375
Hong Kong	0.000***	0.995	0.000***	0.308
Hungary	0.000***	0.222	0.062*	0.329
Iceland	0.000***	0.355	0.000***	0.501
India	0.000***	0.003***	0.067*	0.338
Indonesia	0.000***	0.031**	0.000***	0.954
Ireland	0.000***	0.001***	0.000***	0.574
Italy	0.000***	0.489	0.000***	0.042**
Japan	0.000***	0.012**	0.000***	0.418
South Korea	0.000***	0.098*	0.000***	0.620
Latvia	0.125	0.004***	0.002***	0.284
Lithuania	0.000***	0.001***	0.000***	0.617
Luxembourg	0.000***	0.394	0.000***	0.991
Malta	0.002***	0.001***	0.000***	0.844
Mexico	0.000***	0.000***	0.000***	0.472
Netherlands	0.000***	0.021**	0.019**	0.272
New Zealand	0.000***	0.537	0.000***	0.055*
N. Macedonia	0.000***	0.105	0.000***	0.001***
Norway	0.000***	0.888	0.000***	0.517
Poland	0.000***	0.004***	0.002***	0.014**
Portugal	0.000***	0.987	0.018**	0.105
Romania	0.671	0.007***	0.012**	0.014**
Russia	0.000***	0.187	0.000***	0.479

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (15) and previous for definitions. Cells in gray show encompassed models: They do not explain variation of the combination model, while the competing model does. Cutoff day for PLS = 30.

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 14: Encompassing tests of PLS and benchmark models for unilateral, real exports (II/II)

	h=0		h=1	
	Wald p-Value $H_0^{PLS} : \lambda_1 = 0$	Wald p-Value $H_0^{ARX-B.} : \lambda_2 = 0$	Wald p-Value $H_0^{PLS} : \lambda_1 = 0$	p-Value $H_0^{ARX-B.} : \lambda_2 = 0$
Advanced Asia	0.000***	0.279	0.000***	0.796
Advanced econ.	0.000***	0.000***	0.374	0.613
CIS	0.000***	0.060*	0.000***	0.212
EA 19 extra	0.003***	0.000***	0.065*	0.827
EA 19 intra	0.000***	0.000***	0.677	0.959
EU 27 extra	0.001***	0.000***	0.112	0.963
EU 27 intra	0.000***	0.000***	0.571	0.943
EU candidates	0.000***	0.007***	0.000***	0.147
EU27	0.000***	0.000***	0.279	0.900
EURO	0.000***	0.000***	0.310	0.596
Em. Asia ex CN	0.000***	0.000***	0.001***	0.733
Em. econ. ex CN	0.001***	0.000***	0.015**	0.869
Emerging Asia	0.000***	0.962	0.000***	0.804
Emerging econ.	0.000***	0.209	0.000***	0.740
G20 countries	0.000***	0.000***	0.013**	0.317
Latin America	0.000***	0.003***	0.000***	0.347
Middle-E./ Afr.	0.000***	0.113	0.000***	0.628
Middle-E./ N-Afr.	0.000***	0.327	0.008***	0.169
Other Asia	0.000***	0.011**	0.000***	0.881
Other CIS	0.000***	0.171	0.000***	0.178
Other Latin Am.	0.000***	0.016**	0.000***	0.320
Saudi Arabia	0.015**	0.027**	0.008***	0.247
Singapore	0.000***	0.094*	0.000***	0.958
Slovak Republic	0.000***	0.137	0.000***	0.033**
Slovenia	0.000***	0.093*	0.063*	0.427
South Africa	0.000***	0.431	0.009***	0.565
Spain	0.001***	0.158	0.111	0.060*
Sub-Sah. Africa	0.000***	0.730	0.002***	0.457
Sweden	0.000***	0.001***	0.000***	0.299
Switzerland	0.003***	0.006***	0.000***	0.465
Taiwan	0.012**	0.695	0.020**	0.378
Turkey	0.000***	0.610	0.000***	0.664
UK	0.000***	0.361	0.000***	0.511
USA	0.001***	0.000***	0.022**	0.625
World ex EA19	0.000***	0.001***	0.000***	0.997
World ex EU27	0.001***	0.005***	0.000***	0.852
World	0.000***	0.000***	0.059*	0.933

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (15) and previous for definitions. Cells in gray show encompassed models: They do not explain variation of the combination model, while the competing model does. EA refers to euro area, Em. econ to emerging economies, ex CN to ex China, ex EA19 to ex euro area, Middle-E./ (N.) Afr. to Middle-East and (North) Africa, and Sub-Sah. Afr. designates Sub-Saharan Africa. Cutoff day for PLS = 30.

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01



## A.2 Cutoff After 15 Days

Table 15: Forecast performance of PLS model for unilateral, real imports

	h=0			h=1		
	RMSFE PLS, Abs.	rel. RMSFE Std. Dev.	rel. RMSFE Benchmark	RMSFE PLS, Abs.	rel. RMSFE Std. Dev.	rel. RMSFE Benchmark
Brazil	0.067	0.758***	0.731**	0.077	0.826***	0.835**
China	0.039	0.745***	0.609**	0.042	0.794**	0.786**
EU	0.039	0.942	1.225	0.038	0.919	0.830
France	0.061	0.945	0.914	0.056	0.861	0.841
Germany	0.031	0.893*	1.029	0.032	0.941	0.860
Japan	0.033	0.864**	0.810*	0.035	0.925	0.959
Russia	0.035	0.931***	0.782**	0.033	0.866	0.830
Sub-S. Afr.	0.050	0.945	0.896	0.046	0.875*	0.805
UK	0.065	0.882**	0.869***	0.060	0.813*	0.819*
USA	0.027	0.893	1.083	0.028	0.930	0.891
World	0.027	0.863**	1.364	0.029	0.932	0.836

RMSFE refers to root mean squared forecast error. PLS, Abs. denotes the absolute RMSFE of the PLS model in decimal units. Rel. RMSFE Std. Dev and rel. RMSFE Benchmark denote the relative RMSFEs (13) of the PLS model divided by the standard deviation and the RMSFE of the ARX Benchmark model, respectively. A value below 1 indicates lower forecast errors of the PLS model. Relative RMSFE values above one marked in gray. Benchmark models are ARX-models using economic indicators applicable to the country. Significance levels calculated using Diebold-Mariano test. Sub-S. Afr. designates Sub-Saharan Africa. Cutoff day for PLS = 15.

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 16: Encompassing tests of PLS and benchmark models for unilateral, real imports

	h=0		h=1	
	Wald p-Value $H_0^{PLS} : \lambda_1 = 0$	Wald p-Value $H_0^{ARX-B.} : \lambda_2 = 0$	Wald p-Value $H_0^{PLS} : \lambda_1 = 0$	p-Value $H_0^{ARX-B.} : \lambda_2 = 0$
Brazil	0.000***	0.132	0.000***	0.403
China	0.000***	0.798	0.000***	0.425
EU27	0.000***	0.000***	0.000***	0.494
France	0.015**	0.302	0.001***	0.555
Germany	0.000***	0.000***	0.029**	0.803
Japan	0.001***	0.364	0.001***	0.004***
Russia	0.000***	0.879	0.001***	0.458
Sub-S. Africa	0.000***	0.031**	0.002***	0.138
UK	0.000***	0.470	0.000***	0.647
USA	0.003***	0.000***	0.000***	0.488
World	0.001***	0.000***	0.000***	0.732

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (15) and previous for definitions. Cells in gray show encompassed models: They do not explain variation of the combination model, while the competing model does. Sub-S. Afr. designates Sub-Saharan Africa. Cutoff day for PLS = 15.

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01



Table 17: Forecast performance of PLS model for unilateral, real exports

	h=0			h=1		
	RMSFE PLS, Abs.	rel. RMSFE Std. Dev.	rel. RMSFE Benchmark	RMSFE PLS, Abs.	rel. RMSFE Std. Dev.	rel. RMSFE Benchmark
Brazil	0.065	0.956**	0.882	0.064	0.936***	0.932***
China	0.054	0.809	0.799	0.066	0.977***	0.975
EU	0.047	0.946	1.237	0.047	0.938	0.853
France	0.064	0.942*	0.868	0.061	0.903	0.890
Germany	0.048	0.935	0.945	0.049	0.944	0.880
Japan	0.042	0.975	1.207	0.040	0.918***	0.934
Russia	0.052	0.825***	0.713***	0.050	0.794**	0.727**
Sub-S. Afr.	0.138	0.940	0.915	0.099	0.672	0.581
UK	0.054	0.815*	0.766**	0.064	0.953*	0.904**
USA	0.049	0.905**	1.744	0.051	0.951	0.876***
World	0.028	0.865	1.544	0.030	0.912*	0.849

RMSFE refers to root mean squared forecast error. PLS, Abs. denotes the absolute RMSFE of the PLS model in decimal units. Rel. RMSFE Std. Dev and rel. RMSFE Benchmark denote the relative RMSFEs (13) of the PLS model divided by the standard deviation and the RMSFE of the ARX Benchmark model, respectively. A value below 1 indicates lower forecast errors of the PLS model. Relative RMSFE values above one marked in gray. Benchmark models are ARX-models using economic indicators applicable to the country. Significance levels calculated using Diebold-Mariano test. Sub-S. Afr. designates Sub-Saharan Africa. Cutoff day for PLS = 15.

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 18: Encompassing tests of PLS and benchmark models for unilateral, real exports

	h=0		h=1	
	Wald p-Value $H_0^{PLS} : \lambda_1 = 0$	Wald p-Value $H_0^{ARX-B.} : \lambda_2 = 0$	Wald p-Value $H_0^{PLS} : \lambda_1 = 0$	Wald p-Value $H_0^{ARX-B.} : \lambda_2 = 0$
Brazil	0.005***	0.833	0.000***	0.459
China	0.000***	0.495	0.008***	0.073*
EU27	0.013**	0.000***	0.046**	0.983
France	0.004***	0.710	0.004***	0.305
Germany	0.005***	0.207	0.023**	0.709
Japan	0.187	0.000***	0.000***	0.963
Russia	0.000***	0.086*	0.000***	0.403
Sub-S. Africa	0.041**	0.483	0.000***	0.013**
UK	0.000***	0.439	0.000***	0.612
USA	0.008***	0.000***	0.000***	0.787
World	0.000***	0.000***	0.001***	0.624

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (15) and previous for definitions. Cells in gray show encompassed models: They do not explain variation of the combination model, while the competing model does. Sub-S. Afr. designates Sub-Saharan Africa. Cutoff day for PLS = 15.

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

### A.3 Nominal Trade

Table 19: Forecast performance of PLS model for unilateral, nominal imports

	h=0			h=1		
	RMSFE PLS, Abs.	rel. RMSFE Std. Dev.	rel. RMSFE Benchmark	RMSFE PLS, Abs.	rel. RMSFE Std. Dev.	rel. RMSFE Benchmark
Brazil	0.067	0.752**	0.752**	0.063	0.699***	0.697***
China	0.031	0.772	0.646	0.033	0.821	0.812
EU27	0.031	0.646*	0.679	0.043	0.928	0.830
France	0.047	0.669*	0.803	0.054	0.786	0.786
Germany	0.026	0.637**	0.764	0.033	0.826**	0.740
Japan	0.036	0.820*	0.631	0.030	0.710***	0.717***
Russia	0.025	0.752**	0.601**	0.028	0.853	0.771
UK	0.049	0.718***	0.716***	0.054	0.787***	0.789***
USA	0.024	0.695*	0.949	0.026	0.759**	0.755*

RMSFE refers to root mean squared forecast error. PLS, Abs. denotes the absolute RMSFE of the PLS model in decimal units. Rel. RMSFE Std. Dev and rel. RMSFE Benchmark denote the relative RMSFEs (13) of the PLS model divided by the standard deviation and the RMSFE of the ARX Benchmark model, respectively. A value below 1 indicates lower forecast errors of the PLS model. Relative RMSFE values above one marked in gray. Benchmark models are ARX-models using economic indicators applicable to the country. Significance levels calculated using Diebold-Mariano test. Sub-S. Afr. designates Sub-Saharan Africa. Cutoff day for PLS = 30.

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 20: Encompassing tests of PLS and benchmark models for unilateral, nominal imports

	h=0		h=1	
	Wald p-Value $H_0^{PLS} : \lambda_1 = 0$	Wald p-Value $H_0^{ARX-B.} : \lambda_2 = 0$	Wald p-Value $H_0^{PLS} : \lambda_1 = 0$	p-Value $H_0^{ARX-B.} : \lambda_2 = 0$
Brazil	0.000***	0.076*	0.000***	0.167
China	0.000***	0.577	0.000***	0.407
EU27	0.000***	0.003***	0.003***	0.058*
France	0.000***	0.006***	0.000***	0.008***
Germany	0.000***	0.036**	0.000***	0.119
Japan	0.000***	0.102	0.000***	0.039**
Russia	0.000***	0.024**	0.000***	0.537
UK	0.000***	0.654	0.000***	0.689
USA	0.000***	0.000***	0.000***	0.129

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (15) and previous for definitions. Cells in gray show encompassed models: They do not explain variation of the combination model, while the competing model does. Sub-S. Afr. designates Sub-Saharan Africa. Cutoff day for PLS = 30.

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 21: Forecast performance of PLS model for unilateral, nominal exports

	h=0			h=1		
	RMSFE PLS, Abs.	rel. RMSFE Std. Dev.	rel. RMSFE Benchmark	RMSFE PLS, Abs.	rel. RMSFE Std. Dev.	rel. RMSFE Benchmark
Brazil	0.058	0.936	0.943	0.046	0.751***	0.753***
China	0.059	0.696	0.559*	0.065	0.748	0.726
EU27	0.040	0.724*	0.928	0.050	0.921	0.823
France	0.050	0.711*	0.815	0.062	0.893	0.905
Germany	0.040	0.691	0.751*	0.050	0.867	0.744
Japan	0.028	0.645*	0.872	0.037	0.850*	0.881
Russia	0.049	0.744**	0.771*	0.052	0.807*	0.811
UK	0.064	0.770	0.759	0.064	0.780	0.771
USA	0.048	0.874	2.067	0.051	0.925	0.824

RMSFE refers to root mean squared forecast error. PLS, Abs. denotes the absolute RMSFE of the PLS model in decimal units. Rel. RMSFE Std. Dev and rel. RMSFE Benchmark denote the relative RMSFEs (13) of the PLS model divided by the standard deviation and the RMSFE of the ARX Benchmark model, respectively. A value below 1 indicates lower forecast errors of the PLS model. Relative RMSFE values above one marked in gray. Benchmark models are ARX-models using economic indicators applicable to the country. Significance levels calculated using Diebold-Mariano test. Sub-S. Afr. designates Sub-Saharan Africa. Cutoff day for PLS = 30.

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 22: Encompassing tests of PLS and benchmark models for unilateral, nominal exports

	h=0		h=1	
	Wald p-Value $H_0^{PLS} : \lambda_1 = 0$	Wald p-Value $H_0^{ARX-B.} : \lambda_2 = 0$	Wald p-Value $H_0^{PLS} : \lambda_1 = 0$	p-Value $H_0^{ARX-B.} : \lambda_2 = 0$
Brazil	0.008***	0.873	0.000***	0.343
China	0.000***	0.329	0.000***	0.992
EU27	0.000***	0.000***	0.004***	0.118
France	0.000***	0.224	0.001***	0.065*
Germany	0.000***	0.193	0.000***	0.097*
Japan	0.000***	0.007***	0.002***	0.124
Russia	0.000***	0.064*	0.000***	0.092*
UK	0.000***	0.307	0.000***	0.781
USA	0.175	0.000***	0.017**	0.430

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (15) and previous for definitions. Cells in gray show encompassed models: They do not explain variation of the combination model, while the competing model does. Cutoff day for PLS = 30.

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

## A.4 Example for Variable Selection

Table 23: Variables for estimation of US-American imports in current month

Course/Orientation	Description
Southern course	(20 to 30) <sup>o</sup> lat x (-110 to -120) <sup>o</sup> long
Northern course	(20 to 30) <sup>o</sup> lat x (-100 to -90) <sup>o</sup> long
Western course	(0 to 10) <sup>o</sup> lat x (80 to 70) <sup>o</sup> long
Eastern course	(0 to 10) <sup>o</sup> lat x (-10 to 0) <sup>o</sup> long
Arrival	All ports, USA
Departure	All ports, USA
Arrival	Boston, USA
Arrival	Miami, USA
Arrival	Newark, USA
Departure	Newark, USA
Departure	New York City, USA
Arrival	Tacoma, USA
Arrival	San Vicente, Chile
Arrival	Bremen, Germany
Arrival	Kiel, Germany
Arrival	Conakry, Guinea
Arrival	Chennai, India
Departure	Chennai, India
Departure	Haifa, Israel
Departure	Gioia Tauro, Italy
Departure	San Pedro, Ivory Coast
Arrival	Kobe, Japan
Arrival	Minato (Tokyo), Japan
Arrival	Beirut, Lebanon
Arrival	Port Kelang, Malaysia
Departure	Port Kelang, Malaysia
Arrival	Veracruz, Mexico
Departure	Gdynia, Poland
Arrival	Cape Town, South Africa
Departure	Evyap, Turkey

The order of variables does not reflect the variable importance in the final estimation of the PLS.