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

Vincent Stamer†

December 5, 2022

Abstract

Global container ship movements may reliably predict trade flows. This paper provides the methodology to construct forecasts of trade flows from several million container ship positions per year. An expanding window, out-of-sample exercise shows that constructed forecasts for 76 countries and regions outperform benchmark models. This holds true for unilateral and bilateral trade flows, for trade of developed and developing countries, for real and nominal trade, as well as for time periods of economic crisis such as the COVID-19 pandemic. The resulting forecasts of trade flows precede official statistics by several months and may facilitate quantification of supply chain disruptions and trade wars.

Keywords: Forecasting, Trade, 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, the VfS Annual Meeting, IfW-lunchtime seminar, CAU econometrics seminar, WIFO lunchtime seminar, IFW-staff, Serkan Arslanalp, Diego Alejandro Cerdeiro, Nils Jannsen, Andras Komaromi, Stefan Kooths, Yang Liu, Saskia Meuchelböck, Mariia Okuneva and Alexander Sandkamp for valuable advice.

†Kiel University, Kiel Institute for the World Economy; Wilhelm-Seelig-Platz 1, 24118 Kiel; Phone: +49 4318814228; 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, 2022). Even two years after the first outbreak of the novel virus, 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, decision makers in governments and corporations require information that is as recent as possible. Yet, even the most advanced statistical offices publish unilateral trade data with a lag of several weeks and bilateral data often with a lag of months.

This study applies machine learning techniques on container ship movements to fill this gap. Container vessels moved more than 8 trillion USD worth of goods across the seas in 2020 (UNCTAD, 2021), accounting for almost half of global trade by value. Container ship positions are also available in a matter of hours. This study obtains ship information for the time span from January 2015 to August 2022 for the global fleet of container ships. The raw data has two distinct formats: The first set of data is an exhaustive set of all 450 thousand port calls per year, i.e. the stops of container ships at ports to load or unload cargo. In addition, this study processes over 1,7 Million daily container ship positions per year, which can be aggregated by geographic position and course using a clustering algorithm. Overall, the data yields several Thousand time series of container ship activity. Variable selection with LASSO (least absolute shrinkage and selection operator) in combination with a partial least squares algorithm (PLS) use the shipping data as predictor time series for trade flows. This procedure generates accurate forecasts of trade flows one month ahead, which predate official statistics by several months. Expanding window, out of sample tests show that the forecasts significantly outperform and encompass benchmark models for the vast majority imports and exports of 76 countries and regions, as well as bilateral trade flows between major economies. The combination of LASSO and PLS is shown to be superior to related machine learning tools.

The literature on trade flow prediction is relatively young and fragmented. Veenstra and Haralambides (2001) provide an early work for predicting annual commodity flows using a vector autoregression (VAR) model. In a proof-of-concept study, Cerdeiro et al. (2020) elaborate methods to estimate the cargo of ships in metric tons and show that aggregated ship cargo tends to correlate with countries' imports and exports in the same time period. Furthermore, Circlaeys et al. (2017) employ neural networks to validate annual bilateral trade using GDP values, distances and other control variables from the same time period. Lastly, (Kim, 2020) and (Dubovik et al., 2022) use neural networks and deep learning techniques to forecast South Korean exports and total world trade, respectively. This study serves as a combination and advancement of above studies as this work also processes ship positions to approximate raw cargo flows, but then uses machine learning techniques to make forecasts for the current month and one month ahead of headline trade statistics. To the best of the author's knowledge this is the first paper to introduce forecast methods for imports and exports of 76 countries and regions, as well as bilateral trade flows and to compare them in out-of-sample exercises to benchmark models.

In terms of methodology, this paper follows a growing literature using factor models to reduce dimensionality (see Stock and Watson 2002; Bai and Ng 2004; 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. 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 LASSO as introduced by Tibshirani (1996) and a custom algorithm taking into account the magnitude and geographic proximity of container ship ports. 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 combining LASSO with a PLS model, as well as by outlining aggregation methods for geocoded data. Additionally, above studies combine explanatory time series form a variety of sources for a 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 into country-specific model combinations.

Lastly, this study supplements economic applications of the Automatic Identification System (AIS), which records ship positions and other information based on ships' radio signals. For other uses of AIS data in economics, see, for instance, Brancaccio et al. (2017), Heiland et al. (2019), Sandkamp et al. (2022) and Ganapati et al. (2020). Arslanalp et al. (2019) approximate cargo flows 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 from ship positions and show that aggregate cargo flows tend to correlate with countries' trade. This study does not calculate port calls itself, but receives pre-calculated port calls from a maritime intelligence provider. Instead, this work overlays AIS-derived predictor time series with a machine learning algorithm for forecasting. These algorithms not only trace a country's seaborne actual cargo flows, but also exploit any source of correlation arising through common trends of countries, input-output linkages, as well as substitution and complementary relationships that transcend borders.

The following section describes the conversion of AIS data to independent time series. The chapter also validates pre-calculated port calls and outlines the leading indicators used in the benchmark model. Section 3 on methodology describes the partial least squares model and the required time series tests. Results are presented in section 4 and include benchmark comparisons and encompassing tests. The section separately tests forecasts for real unilateral trade flows, bilateral trade and distinguishes between normal time periods and times of economic crisis. Section 5 discusses the robustness of results to measurement error. Various robustness checks show that results hold for the prediction of nominal and real trade and for the limitation of input data to the first half of the month. The discussion section also compares the above methodology against other machine learning methods such as LASSO, the principal component regression (PCR) and regression trees. The last chapter concludes and gives an outlook for work ahead.

2 Data

The data source for container ship locations used in this study is the Automatic Identification System (AIS), a radio system to monitor vessel movements and avoid collisions at sea. Ships broadcast radio signals every few seconds transmitting ship identifiers, course, speed, GPS position and draught[1]. The maritime intelligence firm FleetMon.com collects the signals from terrestrial receiver stations and satellites, interprets the radio signals and calculates from raw signals the port calls - the stop of cargo ships at a port to load or unload cargo. FleetMon.com provides this project with an exhaustive set of all port calls for container ships over 100 meters in length[2], as well as one AIS position per day for these container ships. The proprietary data spans the time frame from January 2015 to August 2021 and has a global coverage. The 5,265 container ships in the sample constitute 99.5% of all ships labeled ”container ship” in an independently assembled reference list from classification organizations including Lloyd’s List, society DNV-GL, American Bureau of Shipping and Nippon Kaiji Kyoka. The following subsections describe the data processing from port calls to time series and from AIS-positions at sea to time series, respectively.

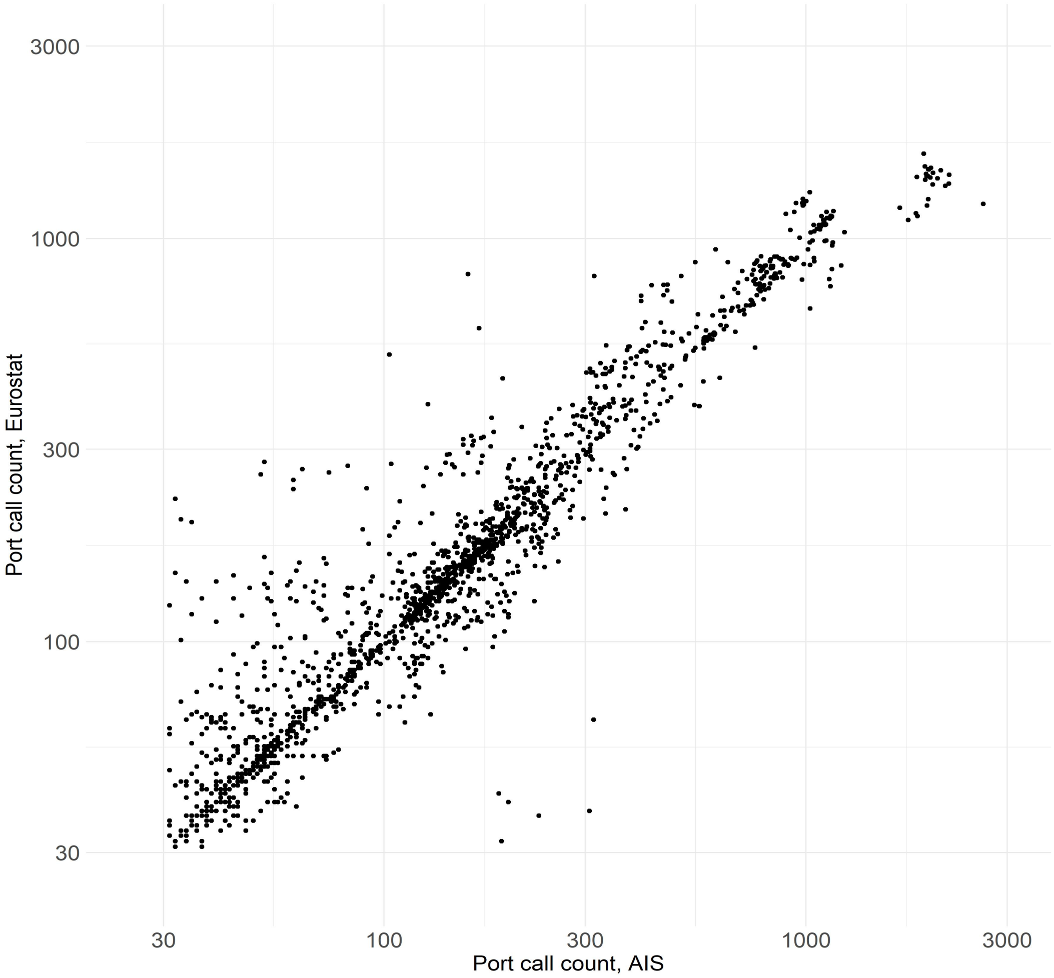
2.1 Predictor time series from port calls

The data on port calls contain approximately 450 Thousand observations per year and cover 1,346 ports in 173 countries. One observation includes ship identifiers, port identifiers, time and draught of arrival, as well as time and draught of departure. The data provider calculates this information based on AIS-positions with a frequency of several seconds and geographic knowledge of port areas. Figure 1 validates the port calls at European ports against quarterly maritime transport statistics published by Eurostat (2017) for container ships of similar size by port.

[1]vertical distance between the waterline and the bottom of the ship’s hull

[2]The smallest container ships, so called feeder ships, which have a carrying capacity of around 300 container only, typically have a length of 130 meters. Hence, the minimum length is not a binding constraint for the vast majority of container ships.

Figure 1: Comparison of number of port calls from Eurostat and AIS



Notes: Quarterly number of port calls from Eurostat for container ships larger than 500 Gross Tonnage at the port level (y-axis) and quarterly number of port calls for container ships larger than 100 meters length from AIS data provider Fleetmon.com (x-axis).

The figure shows great correlation between the two independent data sources. Rotterdam however, the largest port in Europe, does consistently display more port calls derived from the AIS-system than Eurostat. Since only correlation of growth rates in the time series is exploited, but not the absolute number of port calls, consistent mismeasurement has no ramification for the final forecast procedure. Section 5 discusses this further.

In order to measure the container load of ships, the port call data is combined with technical information for ships from MarineTraffic, such as minimum draught (”ballast draught”), maximum draught and capacity measured in Twenty-Foot-Equivalent units (TEU). This technical data in combination with arrival and departure draughts, gives an assessment of the load a ship was carrying when entering and departing a port. A container vessel that approaches a port with a draught close to its minimum value likely carries fewer containers than a vessel with a draught near its maximum, ceteris paribus. 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.

$$TEUload_{it} = TEU_i, max \times \frac{draught_{it} - draught_{i,min}}{draught_{i,max} - draught_{i.min}} \tag{1}$$

Equation 1 translates this into an equation for the TEU load of ship i at time t. This project intentionally does not calculate the net draught change at ports. Since container ships may carry goods towards a port, but may not unload them, Equation 1 may consistently overstate the absolute value of the cargo volume unloaded at a port. Since the PLS model uses growth rates of predictor variables, using gross arrival or departure loads does not have ramifications for final predictions as long as the share of cargo (un)loaded is constant over time. In contrast, calculating the net draught change would lead to a complete loss of information when ships unload and load the same amount of cargo. In this case, the net draught change would be zero despite economically meaningful imports and exports.

To derive time series from this data, the total TEU load of ships entering a port and of ships departing a port are aggregated at the month and port level. This yields two time series per port and more than 2,600 time series overall: One arrival load time series and one departure load time series. To avoid overfitting the factor model with the high number of independent variables, the next step is to reduce further the number of predictor time series. First, time series that have more than 10 monthly observations in the entire time span without any reading are omitted . This lowers the number of port call time series to approximately 1,200. Second, small ports are added to larger ones in geographic proximity. 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 680 and significantly reduces the variance of the smallest ports remaining in the dataset. Furthermore, calculating month-over-month growth rates from above time series derives stationary time series. Using the X-13 procedure of the US Census US Census Bureau (2017) on the predictor 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. 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, the cutoff date in this study is set to be the 30th of the month.

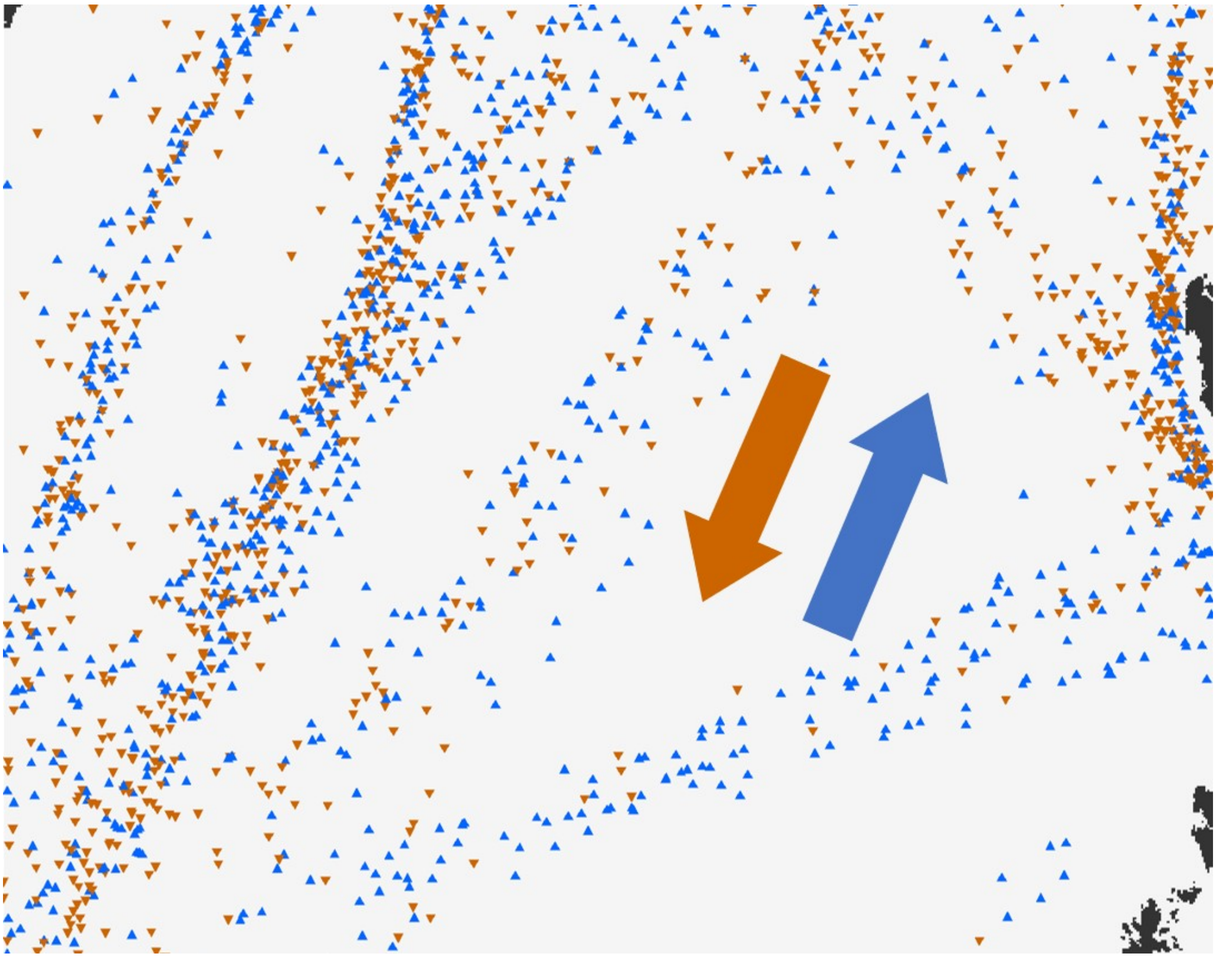
2.2 Predictor time series from AIS-positions

To aggregate the positions at sea to economically meaningful predictor time series, 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 highest number of ship positions are kept. To give an example, Figure 2 shows container ship positions for one such area in the South China Sea during the month of July 2020. The color of the arrows indicate a ship's predominant course. Orange triangles represent ships moving on a Southern course and blue triangles represent ships on a Northern course. Ships going in the opposite direction are likely to carry cargo destined for different countries. A 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.[3] Parallel to the procedure for port calls, single positions are assigned an approximate TEU load by

[3]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.

applying Equation 1. Aggregating the monthly TEU loads of all ships by predominant course yields two time series per area. The resulting 200 series are also seasonally adjusted and can be derived at different cutoff dates during the month.

Figure 2: Illustration of ship data aggregation



The figure shows individual AIS-ship positions in the South China Sea as small triangles. The large arrows illustrate the two predominant courses. For visual purposes approximately 15% of the image are cropped at both the top and bottom.

2.3 Other Data

Trade data for the dependent time stems from CPB (2022) for price adjusted, unilateral trade flows, the WTO (2022) for nominal, unilateral trade flows and IMF (2021) for monthly bilateral trade. The ARX benchmark models are calculated using the following competing leading indicators: The seaborne arrival and departure cargo loads by country are obtained on the IMF website and follow (Cerdeiro et al., 2020). The cargo loads are available beginning with June 2015. European economic sentiment indicators (ESI) and industrial confidence indicators (COF) by country originate from the European Commission (2020). These indicators are matched by country and month to benchmark models. For example, the benchmark model for imports of France includes monthly economic sentiment and industrial confidence indicators for France and the aggregate arriving seaborne cargo load of French ports. The US industrial production (US IP) from the Federal Reserve Bank of St. Louis (2020) and the global supply chain pressure index (GSCPI) from the Federal Reserve Bank of New York (2022) are assigned to every country and region. A port throughput index for the Northrange ports (Döhrn and Maatsch, 2012) from RWI (2020) is assigned to the trade flows of Belgium and the Netherlands. Table 42 in the Appendix lists the leading indicators by country. The benchmark models described below use the seasonally adjusted versions of the indicators if provided and utilize a time series as long as possible.

3 Methodology

3.1 Shrinkage and Partial Least Squares

The aim of this subsection is to outline the methods to reduce dimensionality of the predictor space and calculate forecasts using the partial least squares model. 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 gain popularity because they reduce dimensionality in the span of the predictor time series in the presence of strong colinearity among predictors (Wold et al., 2001). As the number of predictors (N) and the number of time series observations (T) tend towards 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 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. Breitung and Hansen (2021) find in Monte Carlo simulations that T=10 is sufficient to achieve reasonably unbiased estimates for PCR. However, Rönkkö (2014) caution that as N grows the likelihood of chance correlations rises.

This has several 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 smallest frame of the expanding window is chosen reasonably large to avoid bias. In practice, the expanding window analysis begins at half the time periods available so that T is greater or equal to 46. 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 data as outlined in the section above. Additionally, predictor time series are pre-selected using the least absolute shrinkage and selection operator (LASSO) introduced by Tibshirani (1996). It is defined as

$$\hat{\beta} = argmin \left[\sum_t \left(y_t - \sum_j \beta_j x_{tj} \right)^2 + \lambda \sum_j |\beta_j| \right] \tag{2}$$

where penalization for additional coefficients is calculated by the sum of the absolute coefficients multiplied by the shrinkage intensity λ. Beginning with a value of $\lambda = 1$, the parameter is iteratively adjusted so that ultimately 35 to 45 coefficients have a non-zero value. The predictor time series corresponding to non-zero coefficient values are selected. Overall, this reduces the number of predictor time series N below the minimum value of T=46 in the expanding window analysis. The procedure is repeated for every dependent time series and forecast horizon. Note that in the following partial least squares estimation procedure, predictor time series can also receive a zero weight. Hence, even if the lower bound of 35 is still too high for a regression that minimizes forecast errors, the partial least squares approach will not be affected. Table 41 shows pre-selected predictor time series for US-American import nowcasts. Among the 36 predictor time series selected, the list contains several sea areas along the North American coasts, as well as the ports of Miami, New York, Jackson and an aggregate of all US ports. Also several areas and ports in the North Sea, as well as in proximity to the Panama Canal are selected.

The following summarizes the partial least squares process: Factor models such as PLS rely on the idea that a small number of latent factors f_{kt} capture variation for all explanatory time series x_{jt}. The number of latent factors K can be much smaller than the number of predictors J. Factor models approximate these latent factors and use them as components in a model for the dependent time series y_{t+h}.

$$x_{jt} = \sum_{k=1}^{K} \delta_{jk} f_{kt} + \varepsilon_{jt} \qquad \forall x_t \tag{3}$$

$$y_{t+h} = \alpha + \sum_{k=1}^{K} \beta_k \widehat{f}_{kt} + \eta_{t+h} \tag{4}$$

where x_{jt} are predictor time series, f_{kt} refer to latent variables (also known as "factors" or "components"), h is the forecast horizon and ε_{jt} and η_{t+h} are error terms. The PLS algorithm estimates the so called factor loadings δ_{jk}. Ordinary least squares provides estimates of coefficients β. The time subscript, t, is dropped for the following description of components for ease of notation.

An individual component $f_k,\ for\ k = 1, ..., K,$ can be approximated from the J centered explanatory variables $v_{jk},\ for\ j = 1, ..., J$, using

$$\widehat{f}_k = \sum_{j=1}^{J} w_{jk} \phi_{jk} v_{jk} \tag{5}$$

$$v_{jk} = x_{jk} - \bar{\mathbf{x}}_{jk} \tag{6}$$

One derives the component and explanatory variable specific regression coefficients ϕ_{jk} from individually regressing the centered dependent variable u_k on each of the explanatory variables directly:

$$u_k = y_k - \bar{\mathbf{y}}_{\mathbf{k}} \tag{7}$$

$$u_k = \sum_{j=1}^{J} \phi_{jk} v_{jk} \tag{8}$$

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} follows Fuentes et al. (2015) and is set to maximize the covariance between dependent and the independent time series.

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. Cross-validation determines the optimal number of components K for the specific dependent time series and forecast horizon 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 described above, PLS calculates expanding window, direct step 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 $\widehat{y_{t+h}^{PLS}}$. The PLS forecasts are compared to two benchmark models: A naive model set to the mean growth rate and an ARX model with distributed lags up until $p = 1$. The Akaike information criterion determines the number of lags and indicators used in the ARX model. As the PLS inherently forecasts in a direct-step fashion, the ARX fitting follows accordingly:

$$y_{t+h} = \bar{y} \tag{9}$$

$$y_{t+h} = \alpha + \sum_{i=1}^{p} \beta_i \, y_{t-i} + \sum_{i=0}^{p} \sum_{j=0}^{n} \gamma_{ji} \, Indicator_{j,t-i} + \varepsilon_{t+h} \tag{10}$$

where y_{t+h} denotes the month-over-month growth rate of the seasonally adjusted time series from time period $t+h-1$ to $t+h$. The mean $\bar{y}$, as well as intercept α, coefficients β_i and γ_{ji} are estimated in an expanding window method in the same fashion as the actual shipping indicator. The selection of indicators j is summarized in Table 42. The short lag structure is chosen to reflect the high number of predictors relative to the number of observations. Trade flows of Belgium and the Netherlands, for example, are assigned six indicators totaling 12 predictor time series for $p = 1$. To avoid overfitting the benchmark OLS model, this study maintains a ratio of five observations per predictor time series. Hence, the expanding window for ARX benchmark models begins with a frame of 60 observations. As a result, increasing the number of lags further is not feasible. Indicators for bilateral trade flows are a union of the indicators for the two trading partners involved. To reflect the estimation of growth rates, the level data of seaborne cargo flows, US IP and port throughput are also converted to monthly growth rates. Lastly, some of the aforementioned leading indicators are published with a significant time lag. On September 7, 2022, the latest readings for seaborne cargo, US IP and port throughput index were for July. In the main results sections, these indicators are lagged by one month. Robustness checks show that setting $p = 2$ or using indicators without lags does not materially alter the results.

A standard root mean squared forecast error (RMSFE) calculation evaluates the difference between the out-of-sample forecasts and the dependent time series (forecast error, FE) specific to the forecast horizon:

$$RMSFE_h = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(FE_{t+h})^2} \tag{11}$$

A relative RMSFE (rRMSFE) divides the forecast error of the PLS model by the benchmark model's error.

$$relative\ RMSFE_h = \frac{RMSFE_h^{PLS}}{RMSFE_h^{ARX}} \tag{12}$$

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 benchmark models.

Figure 3: Comparison of forecast errors of PLS shipping model against ARX benchmark, h=0

Figure 3 previews results of forecasting imports of individual countries. 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 above tests may show that for some target time series, the benchmark model outperforms the PLS model. The PLS model may nevertheless hold predictive power for the target time series. To test whether that is the case, this paper constructs an encompassing 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.

$$y_{t+h} = \alpha + \sum_{i=1}^{p} \beta_i \, y_{t-i} + \sum_{i=0}^{p} \sum_{j=0}^{n} \gamma_{ji} \, Indicator_{j,t-i} + \delta \, \widehat{y_{t+h}^{PLS}} + \varepsilon_{t+h} \tag{13}$$

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

$$H_0^{PLS} : \lambda_1 = 0 \tag{14}$$

$$H_0^{ARX} : \lambda_2 = 0 \tag{15}$$

where λ_1 and λ_2 are implicitly defined as

$$\widehat{y_{t+h}^{Encomp}} = \alpha + \lambda_1 \, \widehat{y_{t+h}^{PLS}} + \varepsilon_{t+h} \tag{16}$$

$$\widehat{y_{t+h}^{Encomp}} = \alpha + \lambda_2 \, y_{t+h}^{\widehat{ARX-Bench}} + \varepsilon_{t+h} \tag{17}$$

If the Wald test rejects the null hypothesis $H_0^{PLS} : \lambda_1 = 0$, the PLS is not being encompassed and does hold predictive power beyond competing models. Similarly, if the null hypothesis $H_0^{ARX} : \lambda_2 = 0$ 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 the former model's 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 first is disregarded.

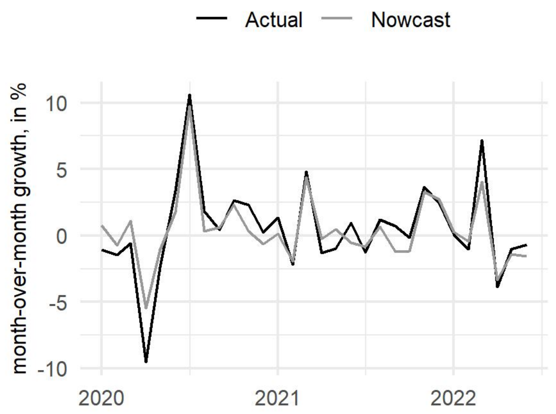
4 Results

4.1 Unilateral Imports

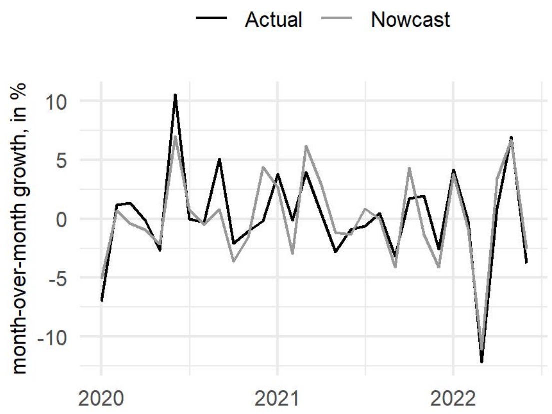
This section reports the estimation results for unilateral imports. Figure 4 begins by plotting the nowcasts against actual import growth for six countries and regions, namely the United States, China, the EU, the world in total, Austria and Sub-Sahara Africa. The black line shows the actual, price and seasonally adjusted monthly growth of imports while the gray line shows the out-of-sample PLS predictions for h=0, i.e. the nowcast. All six panels show a very close fit between dependent variables and the predictions. This holds true both for the landlocked country Austria, as well as the group of developing and emerging countries in Sub-Sahara Africa. Both positive and negative growth are estimated well, although the negative COVID-shock to Austrian and European imports in March and April 2020 is captured only in part. An explanation could be that within the European Union imports via rail and truck were earlier and more severly affected by the ramifications of the pandemic, causing a stronger sudden slump that is reflected in the nowcast based on maritime transport.

The nowcast is available for 76 countries and regions. For a selection of eleven major economies Table 1 quantifies the results more formally. Tables 8 and 9 in the Appendix report all results for all countries and regions. 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

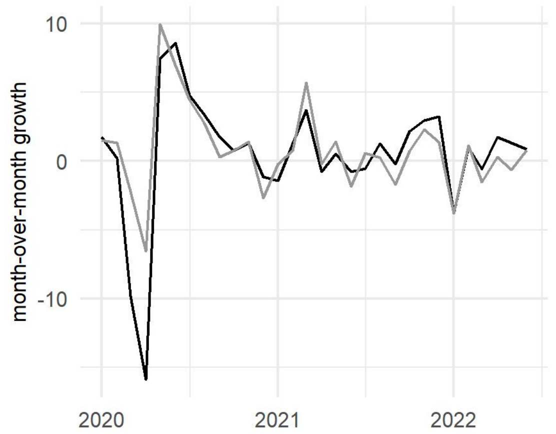
Figure 4: Comparison of nowcast against actual import time series



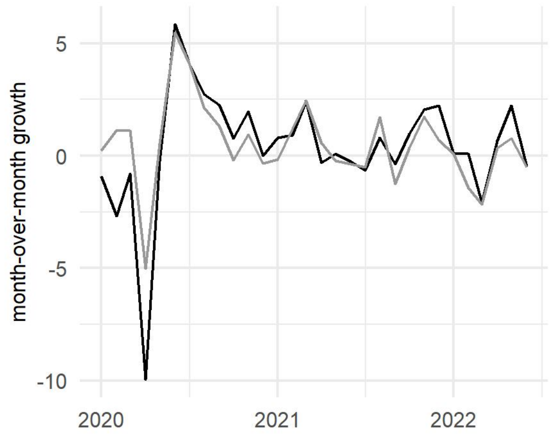
(a) United States



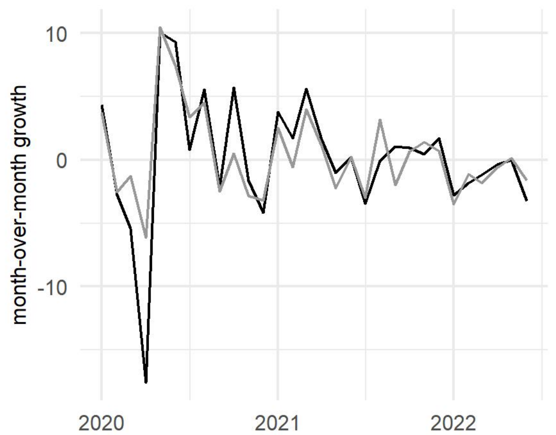
(b) China



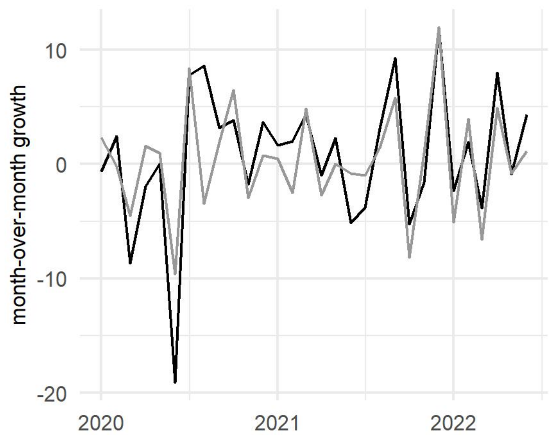
(c) European Union



(d) World



(e) Austria



(f) Sub-Sahara Africa

The figures show the month-over-month growth rate of the actual import time series and the nowcast from the sparse PLS-model.

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 4.8 percentage points. This amounts to approximately 67.0% of the standard deviation of Brazilian imports and to 54.9% of the ARX benchmark model's error. These relative improvements of the PLS model over the other models are highly statistically significant as per the Diebold-Mariano test. The lowest forecast errors are 1.1 pp. and 1.2 pp. for imports of advanced Asia and the world in total, respectively[4]. As comprehensively reported in Tables 8 and 9, the PLS model's forecast errors in $h = 0$ lie below the standard deviation for 53 countries and regions with statistical significance at the $p < 0.05$ cutoff. For the imports of Saudi Arabia and North Macedonia the standard deviation is smaller than the RMSFE of the PLS model, albeit without statistical significance. The PLS model also outperforms the ARX benchmark model with statistical significance for 52 of the countries. Notable exceptions are the group of countries such as the EU, Middle-East /North Africa, Other Asia and other CIS countries.

[4]Import forecasts report similarly low forecast errors for Russian imports. Since the standard deviation of Russian import growth is below that of all other countries and the world in total, one may assume that imports reported by Russia to the CPB are either not credible or adjusted using moving averages. Hence, prediction errors may also be artificially low.

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	0.048	0.670***	0.549***	0.046	0.588***	0.559***
China	0.024	0.508***	0.480***	0.027	0.615***	0.483***
EU	0.021	0.587**	0.717	0.020	0.563	0.506
France	0.031	0.608*	0.324**	0.037	0.771	0.740***
Germany	0.019	0.511**	0.359***	0.021	0.497**	0.386***
Japan	0.020	0.540**	0.353***	0.021	0.581*	0.570***
Russia	0.012	0.411*	0.356**	0.011	0.435**	0.396**
Sub-S. Afr.	0.035	0.679*	0.603*	0.029	0.553**	0.466*
UK	0.039	0.544**	0.442***	0.037	0.466***	0.425**
USA	0.014	0.369**	0.294***	0.014	0.427**	0.380**
World	0.012	0.394*	0.281***	0.015	0.637	0.547*

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 (21) 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.

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Results of the encompassing tests are presented in Table 2, complemented by Tables 10 and 11 in the Appendix. The column labeled ”Wald p-value for H_0^{PLS}: $\lambda_1 = 0$” reports statistical significance with which the null hypothesis is rejected. The table can be interpreted as follows: The Wald test rejects the null hypothesis for the PLS model of Brazilian imports in h=0, namely that the PLS model is irrelevant for the combination model, with a sufficiently small p-value. The PLS model is never encompassed by the ARX benchmark model for any of the countries and regions. In contrast, the new model frequently encompasses the benchmark models. In other words, the PLS model always contributes unique predictive power to the combination model. The only exception is North Macedonia: Neither the PLS-model nor the benchmark model hold predictive power over naive forecasts of North Macedonian imports. Overall, the benchmark model performs very poorly. This points to the fact that trade flows are in general difficult to predict. The discussion in section 5, elaborates on this further and develops robustness checks to improve the performance of the benchmark models.

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***	0.341	0.000***	0.205
China	0.000***	0.470	0.000***	0.125
EU27	0.000***	0.000***	0.000***	0.086*
France	0.000***	0.131	0.000***	0.000***
Germany	0.000***	0.679	0.000***	0.873
Japan	0.000***	0.025**	0.000***	0.362
Russia	0.000***	0.124	0.000***	0.982
Sub-Sah. Africa	0.000***	0.478	0.000***	0.663
UK	0.000***	0.348	0.000***	0.563
USA	0.000***	0.392	0.000***	0.227
World	0.000***	0.464	0.000***	0.437

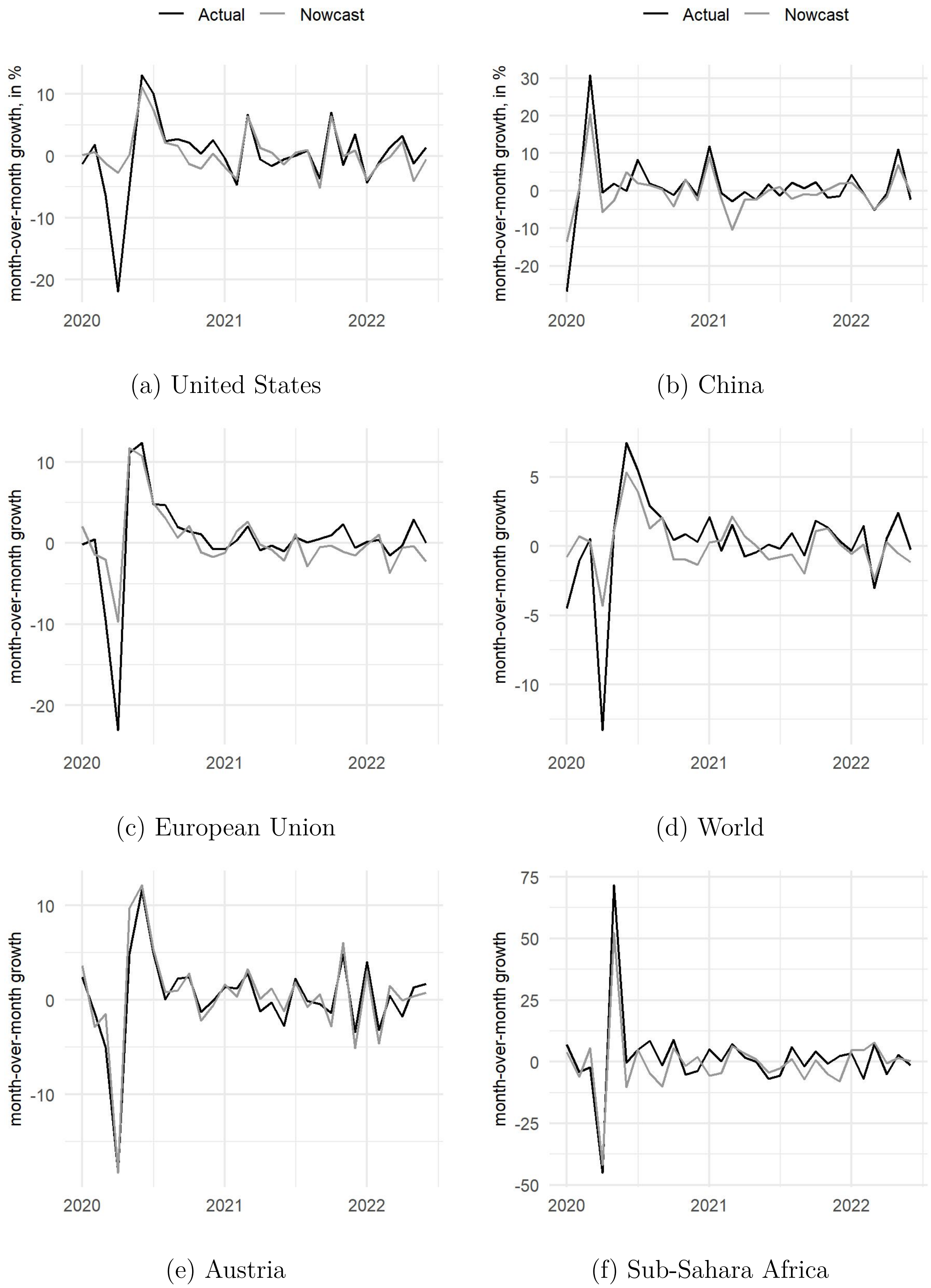
Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (14) and previous for definitions. Cells in gray show encompassed models: They do not explain variation of the combination model, while the competing 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. Figure 5 repeats the visual inspection of forecasts for six sample countries and regions. For the United States, the European Union and the world in total, the PLS model does not track the negative COVID-shock to exports as well as the shock to imports. In contrast, the shipping model works surprisingly well for the landlocked country of Austria. This shows the strength of the model to exploit correlations between German and Italian ports, as well as global trends with Austrian exports.

Figure 5: Comparison of nowcast against actual export time series



(a) United States

(b) China

(c) European Union

(d) World

(e) Austria

(f) Sub-Sahara Africa

The figures show the month-over-month growth rate of the actual export time series and the nowcast from the sparse PLS-model.

Table 3 shows that the results for export forecasts generally match those for imports, although forecast errors for exports are on average 0.5 pp larger. For instance, errors for German imports in $h = 0$ are 1.9 pp, while the errors for German exports are 2.6 pp. Additionally, the forecast errors for exports tend to increase more strongly as one predicts with a forecast horizon of $h = 1$. RMSFEs for German exports, for example, rise from 2.6 pp to 3.6 pp as the forecast period increases. Tables 12 and 13 report the results for all countries and regions. Except for the exports of South Africa, the PLS models' forecast errors lie below the standard deviation and the forecast errors of benchmark models. The PLS-model outperfoms the standard deviation in $h = 0$ with statistical significance at the $p < 0.05$ level for 38 dependent variables and outperforms the benchmark model for 53 dependent variables with statistical significance. The encompassing tests of Tables 4, as well as Tables 14 and 15 show that the PLS model for unilateral exports in the current and subsequent month is never encompassed, but frequently encompasses the benchmark model.

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	0.035	0.442***	0.424***	0.048	0.642***	0.593***
China	0.039	0.755*	0.534***	0.038	0.666**	0.494***
EU27	0.027	0.574*	0.846	0.026	0.553	0.476
France	0.035	0.634	0.440***	0.043	0.533	0.547***
Germany	0.026	0.550*	0.424***	0.036	0.514	0.459***
Japan	0.020	0.443**	0.385***	0.027	0.656**	0.586***
Russia	0.024	0.534	0.405**	0.027	0.609	0.455*
Sub-Sah. Africa	0.063	0.471	0.371	0.064	0.468	0.438
UK	0.041	0.545***	0.534***	0.039	0.450***	0.442***
USA	0.034	0.389**	0.275***	0.041	0.467**	0.396**
World	0.018	0.579*	0.333**	0.017	0.512**	0.488*

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 (21) 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***	0.061*	0.000***	0.865
China	0.000***	0.577	0.000***	0.439
EU27	0.000***	0.000***	0.000***	0.340
France	0.000***	0.293	0.000***	0.080*
Germany	0.000***	0.429	0.000***	0.062*
Japan	0.000***	0.697	0.000***	0.189
Russia	0.000***	0.423	0.000***	0.215
Sub-Sah. Africa	0.000***	0.007***	0.000***	0.107
UK	0.000***	0.380	0.000***	0.640
USA	0.000***	0.986	0.000***	0.184
World	0.000***	0.498	0.000***	0.765

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (14) 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, the algorithms are trained on the bilateral imports of China, the European Union and the United States vis-à-vie their five largest trade partners. Table 5 reports the absolute and relative RMSFE values. The relative RMSFE against the standard deviation are remarkably similar to the performance of the forecasts for unilateral imports. For instance, the RMSFEs of the shipping model for the EU's imports from its top trade partners ranges from 45 to 59% of the standard deviation. This compares to a relative performance of 59% against the standard deviation of imports of the EU in total. Since the baseline standard deviation is much higher for bilateral trade flows, the absolute forecast errors for bilateral trade flows are also higher. As shown in Table 6, the forecasts for bilateral trade flows are never encompassed and in turn frequently encompass the benchmark models.

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	0.054	0.488***	0.419***	0.065	0.614***	0.604***
CN from DE	0.042	0.738*	0.667***	0.039	0.648	0.592***
CN from JP	0.025	0.517***	0.532***	0.031	0.654***	0.660**
CN from KR	0.033	0.602***	0.572***	0.024	0.498***	0.462***
CN from US	0.042	0.684***	0.408	0.030	0.486***	0.473***
EU from CN	0.036	0.544*	0.281	0.035	0.512	0.434*
EU from RU	0.050	0.588**	0.558**	0.060	0.709	0.606
EU from CH.	0.022	0.445***	0.498***	0.025	0.502***	0.409**
EU from UK	0.071	0.504*	0.403**	0.074	0.522	0.496
EU from US	0.036	0.562***	0.596***	0.040	0.624**	0.600**
US from CA	0.035	0.411*	0.263*	0.052	0.630*	0.637
US from CN	0.063	0.787	0.515*	0.047	0.582	0.523
US from DE	0.046	0.636**	0.598**	0.035	0.475**	0.370***
US from JP	0.041	0.584***	0.571***	0.033	0.451***	0.450**
US from MX	0.080	0.627	0.479*	0.071	0.550	0.462

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 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. 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}: $\lambda_2 = 0$	Wald p-Value H_0^{PLS}: λ_1=0	p-Value H_0^{ARX}: $\lambda_2 = 0$
CN from AU	0.000***	0.580	0.000***	0.762
CN from DE	0.000***	0.788	0.000***	0.054*
CN from JP	0.000***	0.131	0.000***	0.969
CN from KR	0.000***	0.294	0.000***	0.165
CN from US	0.000***	0.548	0.000***	0.705
EU from CN	0.000***	0.284	0.000***	0.002***
EU from RU	0.000***	0.001***	0.000***	0.115
EU from CH	0.000***	0.107	0.000***	0.160
EU from UK	0.000***	0.985	0.000***	0.226
EU from US	0.000***	0.023**	0.000***	0.056*
US from CA	0.000***	0.554	0.000***	0.617
US from CN	0.000***	0.691	0.000***	0.003***
US from DE	0.000***	0.882	0.000***	0.936
US from JP	0.000***	0.943	0.000***	0.005***
US from MX	0.000***	0.111	0.000***	0.515

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (14) 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

Largely unexpected shocks such as the COVID-19 crisis make it more difficult to pinpoint the magnitude of disruptions to trade flows. To study the properties of the shipping indicator in times of crisis, this subsection compares the PLS model forecasts during the first half of 2020 and the preceding two years, during which there were not any COVID related lockdowns. A standard AR-model is chosen as the benchmark for two reasons: First, the time span prior to 2018 is not sufficient to train the ARX-model, which requires at least 60 monthly observations for the first out-of-sample estimate and sea cargo by the IMF as a leading indicator begins in mid-2015. Since dependent time series begin in 2014 or earlier, there are ample observations to estimate the AR-model. Second, a direct comparison of autoregression terms with the new shipping forecast highlights the advantages of leading indicators over autoregression terms in general. The standard AR-model follows the following regression function:

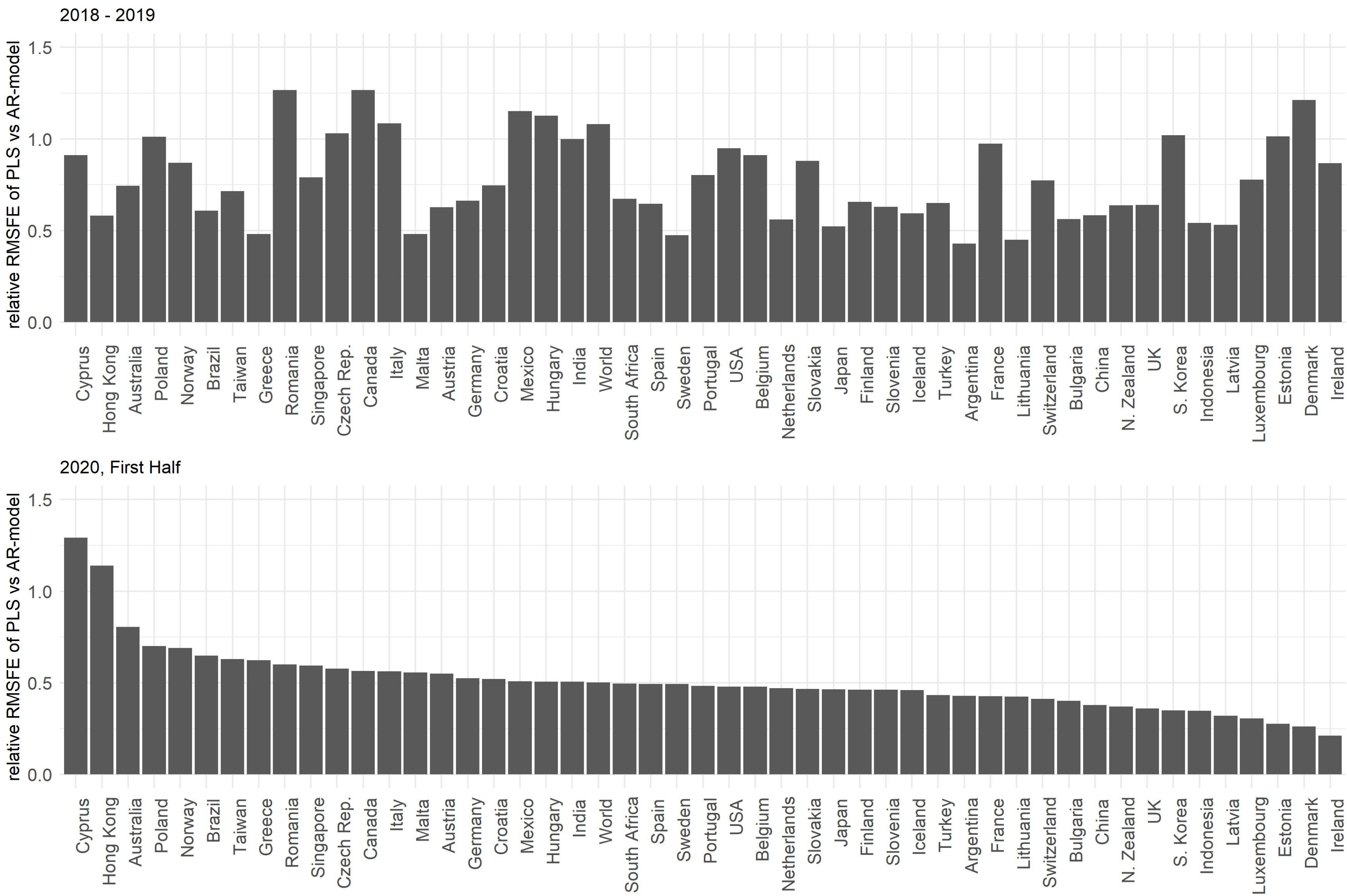
$$y_{t+h} = \alpha + \sum_{i=1}^{3} \beta_i \, y_{t-i} + \varepsilon_{t+h} \tag{18}$$

The AR-model features distributed lags selected by the Akaike information criterion. An expanding window, out-of-sample approach calculates the AR-estimates and a relative RMSFE:

$$relative\, RMSFE_0 = \frac{RMSFE_0^{PLS}}{RMSFE_0^{AR}} \tag{19}$$

Figure 6 summarizes the relative RMSFE of the PLS-model relative to AR-models for all individual country observations and the world in total. In the two years preceding the COVID-19 crisis, the relative RMSFE ranges approximately between 1.25 and 0.5 indicating that for several countries the AR-model produces smaller forecast errors than the PLS-model. In contrast, in the first half of 2020, relative RMSFE averages approximately 0.5 for all countries. This shows that especially in times of economic turmoil the PLS-model is more reliable than AR-models.

Figure 6: Relative RMSFE of new shipping indicator (PLS) vs. AR-model before and during COVID-19 crisis



5 Discussion and Robustness

5.1 Effects of measurement error and structural breaks

Data preparation of the AIS data takes an important part in this study. Hence, this subsection discusses what effect potential measurement errors in the shipping data have on the final forecasts. The first type of potential measurement error concerns an inaccurate approximation of the level of the AIS-derived time series. One may over- or understate the level of economic activity at ports or in sea areas if, for instance, ships are incorrectly classified as “container ships”. The maritime intelligence firm, which provides the AIS data for this study, may also define specific ports differently than other sources. For instance, what is commonly referred to as the “Port of Singapore” is in fact a group of ports with over a dozen terminals. Such variance in classifications may explain why the port of Rotterdam has more AIS port calls per month than recorded by Eurostat in Figure 1. Other sources of measurement error at the level of cargo loads include the salinity and temperature of the sea water and their distorting effect on the ships’ draught or the transport of containers into a port without unloading them. As long as the distortions are constant over time, using monthly growth rates will mask the measurement errors and final forecasts will not change.

A second form of potential measurement error concerns the growth rate of predictor time series. For example, an underestimation of the ballast draught, i.e. minimum draught, or a non-linear relationship between the draught and the cargo volume due to the form of a ship’s hull could lead to an attenuated growth rate in comparison to the actual growth rate of cargo carried. If such measurement error of the growth rate is constant over time, the PLS model will correct this by adjusting the variable regression coefficient ϕ in Equation 5 according to the observed correlation between independent and dependent time series. A similar argument holds for other related trends that change the correlation between dependent and independent time series: Transshipment of cargo implies that not all containers discharged at a port are destined to be imports to the port’s country. Instead, a second vessel likely carries them to a different country. The port of Singapore constitutes a well known example or this. Similarly, many containers unloaded at the port of Rotterdam in the Netherlands are transported by trucks to Germany. These phenomena do not invalidate the PLS forecasts: Both the LASSO and PLS system exploit the empirically observed correlation in the training phase between port cargo flows and country trade flows and adjust the forecast model accordingly.

A more serious threat to the forecasting models are structural breaks in the data, which can be the result of measurement errors or natural occurrences. A one-time change of satellite coverage or manipulation of AIS data could induce such as structural break in the level of the predictor time series. Due to the conversion of time series to growth rates, this structural break only causes the growth rate to display one aberrant observation. If this break occurs in the training phase, it will diminish the correlation between independent and dependent time series and as a result the time series may not be selected for the final forecast model. If this structural break occurs in the testing phase, however, and the predictor time series was selected for the final model, one out-of-sample estimate will be biased. Note that the forecast errors will include this in the error calculation. A similar argument holds for structural breaks in the growth rate of the time series: A break in the testing phase will likely lead to the exclusion of the predictor from the variable selection. A break in the testing phase will increase the measured forecast error. Since the PLS model is re-estimated at every step of the expanding window analysis, even this structural break will impact final forecasts for a very limited number of observations only. As the training window expands, the structural break will move from the testing to the training phase and the PLS model will assign a weight of zero to the predictor time series with a structural break.

5.2 Robustness to time of forecast

To test the robustness of the forecast to different time cutoff, this subsection limits the most recent AIS observations to the 15th day of a month. While previous results rely on observed vessel positions and port calls up to the 30th day of the month, the cutoff date can be set to any day of the month. For example, one may aggregate ship positions in the time period from September 15 to October 15 and assign the observation to the month of October both for training and testing phases of the algorithms. This allows publication of the PLS nowcasts for October in the second half of October. As Tables 16 and 18 show, the absolute forecast errors for the sample economies rise slightly. The forecast errors for imports rise from 1.4 pp (30 day cutoff, Table 16) to 2.3 pp (15 day cutoff) for the United States and even fall from 3.1 pp to 2.9 pp for France. The rise of forecast errors for the other highly developed countries is typically below 0.5 pp. The forecast errors for exports increase to a slightly higher extent. While relative RMSFE values also remain below 1 and signal that the PLS model continues to outperform benchmark models, there are weaknesses of the export forecasts. The benchmark produces lower forecast errors for nowcasts of EU exports and both the standard deviation and benchmark RMSFEs are lower than forecast errors for French exports. These results do not carry statistical significance and the PLS model still contributes unique predictive power to the forecasts of European exports according to the rejection of the null hypothesis $H_0^{PLS} : \lambda_1 = 0$ in Table 19. The PLS forecasts for French exports are encompassed by the benchmark model.

5.3 Robustness to price adjustment of trade flows

The second robustness check concerns the price adjustment of the dependent variable. The previously estimated unilateral trade flows are seasonally and price adjusted unilateral trade flows from CPB (2022). In contrast, this subsection repeats the forecasting exercise for unilateral, nominal trade flows from the WTO (2022). The dependent variables are also seasonally adjusted using the X-13 ARIMA method. Tables 20 and 21 show the benchmark comparisons and encompassing tests for nominal imports. Absolute forecast errors do not change for Brazil, France and Germany, fall for the EU and the United Kingdom and increase by approximately 0.6 pp for China, Japan and the United States. Although the absolute estimation errors remain constant for French and German imports, the relative RMSFE against the benchmarks increases. This indicates that the benchmark model performs better for nominal trade flows. Tables 22 and 23 show the respective results for exports. Overall, estimation errors tend to rise for nowcasts of nominal exports as compared to nowcasts of real exports. The increase in absolute RMSFEs ranges from zero for the European Union to 1.6 pp for China. As a result, relative RMSFEs for the comparison with benchmark models increase and surpass the critical value of one for the United States. The value is not statistically significant, however, and the benchmark model does not encompass the PLS model's nowcast.

5.4 Robustness to benchmark model specifications

This study prominently features the ARX model as a benchmark for the performance of the PLS model and the AIS derived predictor time series. Somewhat surprisingly, the ARX model performs poorly for many countries and the historic average growth rate is frequently a better predictor of future growth rates than the ARX model. One explanation for this is that trade flows, month-over-month growth time series with high variance in particular, are difficult to predict. Another reason is given above: The benchmark models perform better for nominal trade flows. Both explanations do not diminish the credibility of the benchmark test in above analyses. One would conclude that the sparse PLS model produces reasonable forecasts *despite* the fact that dependent time series are difficult to predict and that the shipping data predicts real trade flows relatively better than nominal trade flows.

A more concerning problem would be if the ARX model is incorrectly specified and PLS model results appear reasonable only in comparison to a misspecified model. The following four robustness checks test the sensitivity of the results to other specifications of the benchmark model. The first robustness check expands the lags of the benchmark model. In above specifications summarized by Equation 10, the distributed lags of the benchmark model include the current time period, $t = 0$, as well as the previous time period $t = -1$. Adding further lags implies a trade off between potential predictive power of past lags against overfitting the OLS model with too many independent time series. This robustness check repeats above analyses but includes an additional lag in the ARX model, i.e. observations in $t = -2$. The Akaike information criterion determines the distribution of lags. Tables 24 and 25 in the Appendix display the results for unilateral, real imports and Tables 26 and 27 show the results for unilateral, real exports. The relative RMSFEs in the columns labeled ”rel RMSFE Benchmark” increase over previous results. Hence, an additional lag does not improve the benchmark model.

Figure 6 shows that the autoregression terms lead to relatively poor predictions during COVID-related disruptions. As a result, this robustness check tests if the benchmark model improves when omitting the autoregression terms. The regression formula for the benchmark model, therefore, is:

$$y_{t+h} = \alpha + \sum_{i=0}^{p} \sum_{j=0}^{n} \gamma_{ji}\, Indicator_{j,t-i} + \varepsilon_{t+h} \tag{20}$$

Tables 28 and 29 show the results for imports, while Tables 30 and 31 display the results for exports. Indeed, the relative RMSFE values for both imports and exports generally increase. This points to the fact that the autoregression terms and structural breaks in their correlation with the dependent variable deteriorate the benchmark model. The relative forecast errors are nevertheless still below one and the PLS model continues to outperform the benchmark model.

Several of the leading indicators are published with a significant time lag. For instance, the maritime cargo loads calculated by Cerdeiro et al. (2020), as well as the RWI/ISL throughput index for ports are published with a time lag of approximately a month. Hence, the researcher would have to wait until at least early December in order to calculate the ”nowcast” for the preceeding October. The PLS model can instead calculate a nowcast for October at any time in the second half of October or early November. To reflect this time advantage, several competing leading indicators are lagged by one period as outlined in Section 3. While this is a reflection of reality, one can nevertheless calculate a robustness check without this additional time. The robustness check, therefore, simulates how well the ARX model would forecast if all leading indicators were available at the same time as the PLS nowcasts. Tables 32 and 33 show the results for unilateral, real import time series and Tables 34 and 35 show the respective results for unilateral, real export time series.

Again, the relative RMSFEs between PLS and benchmark models increase which points to improved benchmark models. This is, of course, to be expected as the leading indicators are simulated to be available without time lag. The PLS model still outperforms the benchmark, except for European Union exports with forecast horizon $h = 0$. The relative RMSFE is 1.200 indicating that the benchmark produces 20% lower forecast errors than the model. This result is, however, not statistically significant and the PLS model is not encompassed by the competing model.

In a final robustness check, the benchmark model omits the maritime cargo loads calculated by Cerdeiro et al. (2020) and disseminated by the IMF. Due to the focus on maritime cargo the estimates are a natural benchmark for this study. Unfortunately, this data source starts in June 2015 reducing the training period for the ARX model. Several leading indicators such as export expectations (Grimme and Wohlrabe, 2014) or import climate (Grimme et al., 2021) from the ifo Institute (Ifo institute, 2021), port throughput data (Döhrn and Maatsch, 2012), as well as manufacturing order entries from abroad published by Destatis (2021) replace the omitted sea cargo for the forecast of German imports and exports. As a result, the dependent time variables and leading indicator time series are complete starting in 2007. This enlarges the training phase for the ARX model threefold. Tables 36 and 37 report the results for imports, whereas Tables 38 and 39 report the results for exports. In line with previous robustness checks, the relative forecast errors of the PLS model against the benchmark increase. Both for German and US-American imports and exports, the values approximately double. This shows that the longer training period does improve the benchmark forecasts. The relative forecast errors do remain below one, however, which signifies that the PLS model continues to produce lower forecast errors than the benchmark.

5.5 Performance against other machine learning algorithms

The objective of above analyses is to show that the sparse PLS model generates estimates that are useful for forecasting in absolute terms and relative to benchmark models. A secondary question remains whether the sparse PLS model is the best machine learning

(ML) tool to estimate the trade flows with the AIS data. To answer this, LASSO as a stand-alone estimator, a regression tree using cross-validation methods for ”pruning” – see, (Januschowski et al., 2021) for an overview of associated methods –, a principle component regression (PCR) and a PLS system are trained on shipping data with and without shrinkage by LASSO. The results are compared to the PLS model with shrinkage as used above using the relative RMSFE indication:

$$relative\ RMSFE_h = \frac{RMSFE_h^{SparsePLS}}{RMSFE_h^{Other\ ML}} \tag{21}$$

Table 7 reports the results for imports and Appendix Table 40 reports the result for exports. The first column “Sparse PLS, abs. RMSFE” lists the absolute forecast error of the out-of-sample forecasts of the sparse PLS model as used previously. The subsequent columns measure the relative forecast error as compared to the competing ML tools. A first insight is that shrinkage tends to improve the forecasts as the relative RMSFEs of the sparse methods are higher than those without shrinkage. This holds true for the PCR model in particular. Similarly, the relative RMSFE against the PLS model without shrinkage is around 0.5 which indicates that the variable selection by LASSO lowers the forecasting error by half. LASSO reduces the signal to noise ratio which in turn improves the factor models PCR and PLS. While the sparse PLS model outperforms the other machine learning tools, the sparse PCR model produces forecasts with very similar accuracy. This is not surprising as the two factor models are closely related. The biggest difference between the models is that PLS takes into account the covariance between predictor and dependent variables when forming the components, while PCR only considers the covariance between predictors. This difference is, however, masked by LASSO, which also takes into account the correlation between predictor and dependent time series. Of course, the above list of machine learning tools is not exhaustive. Boosting the regression tree with a random forest approach, employing Bayesian estimation methods and applying neural networks may produce different results. Nevertheless, the exercise above shows that shrinkage and factor models may outperform other standard tools.

Table 7: Comparison of various machine learning tools with and without sparcity against the sparse PLS; real imports

	Sparse PLS abs.RMSFE	Lasso rRMSFE	PCR rRMSFE	Reg. Tree rRMSFE	PLS rRMSFE	Sparse Lasso rRMSFE	Sparse PCR rRMSFE	Sparse Tree rRMSFE
h=0								
Brazil	0.047	0.570***	0.596***	0.476***	0.589***	0.631***	0.978	0.541***
China	0.024	0.496***	0.471***	0.433***	0.500***	0.519***	0.994	0.490***
EU27	0.022	0.608**	0.602**	0.591**	0.524	0.623**	1.052	0.637***
France	0.032	0.492*	0.499*	0.473**	0.483	0.512*	0.965	0.529**
Germany	0.018	0.550**	0.505*	0.572**	0.509*	0.554**	0.913	0.498***
Japan	0.020	0.514***	0.506***	0.502**	0.528***	0.534***	0.928	0.498***
Russia	0.012	0.570*	0.559*	0.514	0.492**	0.570*	0.983	0.642
Sub-Sah. Africa	0.035	0.689*	0.696*	0.643**	0.691*	0.686**	0.962	0.655**
UK	0.041	0.562**	0.546**	0.573***	0.536***	0.588**	0.991	0.618***
USA	0.014	0.465**	0.421**	0.436**	0.455***	0.466**	0.818*	0.449**
World	0.012	0.581*	0.577*	0.562*	0.513**	0.581*	1.004	0.540*
h=1								
Brazil	0.049	0.614***	0.615***	0.530***	0.589***	0.682***	0.861**	0.647***
China	0.027	0.560***	0.545**	0.508**	0.566**	0.562***	0.940*	0.554***
EU27	0.018	0.512	0.480	0.455	0.441	0.510	0.774	0.486
France	0.036	0.535	0.519	0.521	0.523	0.563	0.840	0.560
Germany	0.022	0.669**	0.651**	0.643**	0.628**	0.670**	0.916	0.611***
Japan	0.021	0.486*	0.562**	0.473**	0.556*	0.499*	0.952	0.453***
Russia	0.011	0.489*	0.490**	0.504*	0.461***	0.489*	0.862**	0.495**
Sub-Sah. Africa	0.029	0.575**	0.555**	0.577*	0.591**	0.631**	0.969	0.554**
UK	0.040	0.538***	0.544***	0.502***	0.510***	0.585**	0.957	0.513***
USA	0.014	0.475**	0.442**	0.433**	0.432**	0.474**	0.860	0.461***
World	0.015	0.715	0.707	0.672*	0.593	0.715	0.804	0.667**

Note: The column "Sparse PLS, abs. RMSFE" measures the absolute forecast error of the sparse partial least squares model. Columns labeled "rRMSFE" measure the relative RMSFE of the sparse PLS as compared to the noted machine learning tools. "Sparse" always refers to prior shrinkage of the predictor space using LASSO.

6 Conclusion

This paper introduces methods to convert high frequency, geocoded data from the container ship network into predictor time series useful for economic forecasting. Since these methods can derive 880 or more independent time series, the paper combines a least absolute shrinkage and selection operator (LASSO) with a partial least squares model to reduce dimensionality of the predictor space. The model is then used to forecast unilateral and bilateral trade flows for the current time period and one period ahead. Forecasts frequently outperform and encompass benchmark models with competing leading indicators for 76 countries and regions. Extensions show that these results hold for real and nominal trade flows, for developed and developing nations and in times of economic crisis during the COVID-19 pandemic. Factor model such as PLS are well suited for an application to the shipping data and gain predictive power when combined with LASSO. Furthermore, overlaying the raw shipping data with a forecasting model has several advantages: All predictions made by the sparse PLS model can be readily interpreted as monthly growth of national trade flows as opposed to leading indicators that measure the absolute cargo volume arriving or departing ports. The prediction model also exploits sources of correlation between a country and foreign ports such as input-output linkages between countries. 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 such as the economic activity of subnational regions or specific industries.

Declaration of Competing Interest

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Arslanalp, S., Marini, M., and Tumbarello, P. (2019). Big Data on Vessel Traffic: Nowcasting Trade Flows in Real Time. Technical report, IMF Working Papers 2019/275.

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.

Brancaccio, G., Kalouptsidi, M., and Papageorgiou, T. (2017). Geography, search frictions and endogenous trade costs. Technical report, NBER Working Papers 23581.

Breitung, J. and Hansen, P. (2021). Alternative estimation approaches for the factor augmented panel data model with small t. *Empirical Economics*, 60(1):327–351.

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.

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., Mineshima, A., Hansen, N.-J., Zhou, J., and Spector, M. (2022). Supply bottlenecks: Where, why, how much, and what next? Technical report, IMF Working Papers 2022/031.

Cerdeiro, M. D. A., Komaromi, A., Liu, Y., and Saeed, M. (2020). World Seaborne Trade in Real Time: A Proof of Concept for Building AIS-based Nowcasts from Scratch. Technical report, IMF Working Papers 2020/057.

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.

Circlaeys, S., Kanitkar, C., and Kumazawa, D. (2017). Bilateral trade flow prediction. Technical report, Stanford University.

CPB (2022). CPB Netherlands Bureau for Economic Policy Analysis. World Trade Monitor.

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.

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.

Dubovik, A., Elbourne, A., Hendriks, B., Kattenberg, M., et al. (2022). Forecasting world trade using big data and machine learning techniques. Technical report, CPB Netherlands Bureau for Economic Policy Analysis.

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.

Eurostat (2017). International trade by mode of transportation.

Federal Reserve Bank of New York (2022). Global Supply Chain Pressure Index.

Federal Reserve Bank of St. Louis (2020). Board of Governors of the Federal Reserve System (US), Industrial Production.

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.

Ganapati, S., Wong, W. F., and Ziv, O. (2020). Entrepôt: Hubs, scale, and trade costs. Technical report, CESifo Working Paper Series 8199.

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. (2021). Forecasting imports with information from abroad. *Economic Modelling*, 98(C):109–117.

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. Technical report, CEPR Discussion Papers 14193.

Ifo institute (2021). Various leading indicators. Ifo Zeitreihen - Exporterwartungen, Import- and Exportklima.

IMF (2021). International Monetary Fund. DOTS.

Januschowski, T., Wang, Y., Torkkola, K., Erkkilä, T., Hasson, H., and Gasthaus, J. (2021). Forecasting with trees. *International Journal of Forecasting*.

Jurado, K., Ludvigson, S. C., and Ng, S. (2015). Measuring uncertainty. *American Economic Review*, 105(3):1177–1216.

Kim, S. (2020). Macroeconomic and financial market analyses and predictions through deep learning. *Bank of Korea WP*, 18.

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 (2020). RWI/ISL Containerumschlag-Index.

Sandkamp, A., Stamer, V., and Yang, S. (2022). Where has the rum gone? the impact of maritime piracy on trade and transport. *Review of World Economics*, 158:751–778.

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.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society*, 58(1):267–288.

UNCTAD (2021). Review of maritime transport 2021.

US Census Bureau (2017). X-13ARIMA-SEATS Reference Manual.

Veenstra, A. W. and Haralambides, H. E. (2001). Multivariate autoregressive models for forecasting seaborne trade flows. *Transportation Research Part E*, 37(4):311–319.

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.

WTO (2022). World Trade Organization. WTO DATA.

A Appendix

A.1 Full Country Results

Table 8: 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	0.043	0.651***	0.568***	0.036	0.527***	0.497***
Australia	0.027	0.670***	0.495***	0.027	0.568**	0.564**
Austria	0.024	0.562***	0.520***	0.023	0.521*	0.488*
Belgium	0.024	0.528**	0.350***	0.024	0.478***	0.405***
Brazil	0.048	0.670***	0.549***	0.046	0.588***	0.559***
Bulgaria	0.025	0.594***	0.298*	0.031	0.731***	0.667**
Canada	0.037	0.405*	0.297**	0.035	0.401	0.328*
China	0.024	0.508***	0.480***	0.027	0.615***	0.483***
Croatia	0.033	0.583**	0.317*	0.036	0.723	0.436***
Cyprus	0.151	0.813	0.290*	0.117	0.605**	0.475***
Czech Rep.	0.027	0.607**	0.484**	0.025	0.558	0.541
Denmark	0.020	0.622**	0.474**	0.020	0.674*	0.571***
Estonia	0.026	0.731*	0.606*	0.022	0.523**	0.426***
Finland	0.018	0.380***	0.245**	0.019	0.429***	0.347***
France	0.031	0.608*	0.324**	0.037	0.771	0.740***
Germany	0.019	0.511**	0.359***	0.021	0.497**	0.386***
Greece	0.035	0.573***	0.381***	0.038	0.564***	0.535***
Hong Kong, China	0.023	0.360***	0.356***	0.018	0.410**	0.345**
Hungary	0.030	0.609**	0.515**	0.026	0.525*	0.506*
Iceland	0.029	0.499	0.488**	0.025	0.409	0.398*
India	0.038	0.512*	0.341**	0.048	0.761	0.769***
Indonesia	0.055	0.735**	0.537***	0.052	0.710**	0.652***
Ireland	0.052	0.471**	0.290***	0.042	0.366**	0.272*
Italy	0.032	0.745**	0.407***	0.031	0.888	0.691***
Japan	0.020	0.540**	0.353***	0.021	0.581*	0.570***
Korea, Republic of	0.018	0.327***	0.323***	0.017	0.362***	0.345***
Latvia	0.021	0.546***	0.549***	0.027	0.604***	0.563**
Lithuania	0.015	0.517***	0.339***	0.015	0.440***	0.366***
Luxembourg	0.047	0.576**	0.513***	0.054	0.620*	0.597*
Malta	0.107	0.510***	0.558**	0.114	0.468***	0.448**
Mexico	0.042	0.425**	0.413**	0.033	0.337*	0.240*
Netherlands	0.014	0.447***	0.314***	0.015	0.489***	0.337***
New Zealand	0.027	0.400***	0.348***	0.036	0.453***	0.385***
N. Macedonia	0.013	1.833***	0.216**	0.014	1.734***	0.520
Norway	0.038	0.881*	0.932	0.040	0.757	0.738**
Poland	0.030	0.410	0.292**	0.027	0.421	0.351**
Portugal	0.038	0.493**	0.479***	0.040	0.652*	0.543***
Romania	0.045	0.679*	0.752	0.041	0.604	0.591
Russia	0.012	0.411*	0.356**	0.011	0.435**	0.396**

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 (21) 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 9: 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	0.011	0.421***	0.407***	0.013	0.484***	0.435***
Adv. econ.	0.012	0.522**	0.366*	0.018	0.747*	0.720
CIS	0.021	0.566	0.447**	0.022	0.595	0.570*
EA 19 extra	0.017	0.634***	0.420**	0.014	0.526***	0.513***
EA 19 intra	0.031	0.653**	0.485*	0.032	0.667	0.645
EU 27 extra	0.016	0.611***	0.405**	0.013	0.496***	0.485***
EU 27 intra	0.031	0.654*	0.479	0.028	0.583	0.570
EU candidate cntrs	0.035	0.516**	0.504***	0.044	0.650*	0.630*
EU27	0.021	0.587**	0.717	0.020	0.563	0.506
EURO	0.023	0.551**	0.488***	0.022	0.549	0.532***
EM A. exCN	0.026	0.703*	0.620**	0.024	0.679*	0.595*
EM exCN	0.016	0.638**	0.547**	0.017	0.711**	0.644**
EM Asia	0.022	0.673***	0.498*	0.019	0.597***	0.571**
EM	0.015	0.586***	0.472**	0.016	0.654***	0.620***
G20 countr.	0.013	0.554*	0.397*	0.014	0.592*	0.573*
Latin Am.	0.029	0.686***	0.658*	0.026	0.620***	0.593**
M.-E./ Afr.	0.010	0.685	0.441	0.010	0.721*	0.528*
M.-E./ N.Afr.	0.009	0.951	0.571	0.010	0.969	0.620
Other Asia	0.023	0.585**	0.527	0.027	0.680	0.598
Other CIS	0.040	0.530	0.501	0.057	0.750*	0.740*
O. Latin Am.	0.027	0.507***	0.377*	0.032	0.594***	0.546***
Saudi Arab.	0.025	1.006*	0.239	0.026	0.798	0.623***
Singapore	0.019	0.515***	0.440***	0.019	0.593***	0.483***
Slovak Republic	0.039	0.612*	0.537**	0.038	0.584	0.546
Slovenia	0.037	0.421***	0.395***	0.040	0.390***	0.341***
South Africa	0.044	0.599**	0.678*	0.033	0.413**	0.321**
Spain	0.032	0.546**	0.294***	0.030	0.605	0.538**
Sub-S. Afr.	0.035	0.679*	0.603*	0.029	0.553**	0.466*
Sweden	0.015	0.608**	0.389*	0.015	0.486**	0.456***
Switzerland	0.026	0.487**	0.438**	0.030	0.559**	0.528**
Taiwan	0.022	0.543***	0.477***	0.024	0.608***	0.519***
Turkey	0.034	0.468**	0.265*	0.044	0.702*	0.588*
UK	0.039	0.544**	0.442***	0.037	0.466***	0.425**
USA	0.014	0.369**	0.294***	0.014	0.427**	0.380**
World ex EA	0.012	0.549*	0.433*	0.014	0.670*	0.639*
World ex EU	0.010	0.488**	0.399**	0.014	0.695*	0.658
World	0.012	0.394*	0.281***	0.015	0.637	0.547*

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 (21) 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 10: 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***	0.646	0.000***	0.621
Australia	0.000***	0.129	0.000***	0.451
Austria	0.000***	0.372	0.000***	0.117
Belgium	0.000***	0.121	0.000***	0.529
Brazil	0.000***	0.341	0.000***	0.205
Bulgaria	0.000***	0.880	0.000***	0.763
Canada	0.000***	0.706	0.000***	0.609
China	0.000***	0.470	0.000***	0.125
Croatia	0.000***	0.215	0.000***	0.575
Cyprus	0.003***	0.300	0.000***	0.540
Czech Rep.	0.000***	0.997	0.000***	0.072*
Denmark	0.000***	0.918	0.000***	0.678
Estonia	0.000***	0.858	0.000***	0.169
Finland	0.000***	0.792	0.000***	0.771
France	0.000***	0.131	0.000***	0.000***
Germany	0.000***	0.679	0.000***	0.873
Greece	0.000***	0.726	0.000***	0.710
Hong Kong, China	0.000***	0.330	0.000***	0.660
Hungary	0.000***	0.437	0.000***	0.748
Iceland	0.000***	0.401	0.000***	0.225
India	0.000***	0.323	0.000***	0.466
Indonesia	0.000***	0.927	0.000***	0.430
Ireland	0.000***	0.080*	0.000***	0.839
Italy	0.000***	0.266	0.006***	0.207
Japan	0.000***	0.025**	0.000***	0.362
Korea, Republic of	0.000***	0.043**	0.000***	0.063*
Latvia	0.000***	0.184	0.000***	0.031**
Lithuania	0.000***	0.090*	0.000***	0.835
Luxembourg	0.000***	0.410	0.000***	0.054*
Malta	0.000***	0.374	0.000***	0.094*
Mexico	0.000***	0.836	0.000***	0.056*
Netherlands	0.000***	0.654	0.000***	0.581
New Zealand	0.000***	0.086*	0.000***	0.410
N. Macedonia	0.235	0.854	0.703	0.341
Norway	0.003***	0.294	0.000***	0.530
Poland	0.000***	0.397	0.000***	0.361
Portugal	0.000***	0.991	0.000***	0.154
Romania	0.000***	0.000***	0.000***	0.039**
Russia	0.000***	0.124	0.000***	0.982

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (14) 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 11: 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 $H_0^{ARX-B.}$: $\lambda_2 = 0$
Adv. Asia	0.000***	0.875	0.000***	0.228
Adv. econ	0.000***	0.103	0.000***	0.100
CIS	0.000***	0.421	0.000***	0.740
EA 19 extra	0.000***	0.440	0.000***	0.171
EA 19 intra	0.000***	0.069*	0.000***	0.478
EU 27 extra	0.000***	0.262	0.000***	0.636
EU 27 intra	0.000***	0.026**	0.000***	0.201
EU candidates	0.000***	0.348	0.000***	0.985
EU27	0.000***	0.000***	0.000***	0.086*
EURO	0.000***	0.880	0.000***	0.112
Em. Asia ex CN	0.000***	0.297	0.000***	0.999
Em. econ ex CN	0.000***	0.157	0.000***	0.974
EM Asia	0.000***	0.018**	0.000***	0.030**
EM econ.	0.000***	0.672	0.000***	0.136
G20 countries	0.000***	0.335	0.000***	0.171
Latin America	0.000***	0.069*	0.000***	0.005***
Mid.-E./ Afr.	0.000***	0.146	0.000***	0.177
Mid.-E./ N-Afr.	0.000***	0.524	0.000***	0.591
Other Asia	0.000***	0.397	0.000***	0.361
Other CIS	0.000***	0.124	0.000***	0.106
Other Lat. Am.	0.000***	0.975	0.000***	0.001***
Saudi Arab.	0.000***	0.754	0.000***	0.764
Singapore	0.000***	0.709	0.000***	0.031**
Slovak Republic	0.000***	0.866	0.000***	0.552
Slovenia	0.000***	0.357	0.000***	0.253
South Africa	0.000***	0.217	0.000***	0.971
Spain	0.000***	0.021**	0.000***	0.618
Sub-Sah. Afr.	0.000***	0.478	0.000***	0.663
Sweden	0.000***	0.442	0.000***	0.664
Switzerland	0.000***	0.364	0.000***	0.647
Taiwan	0.000***	0.304	0.000***	0.545
Turkey	0.000***	0.586	0.000***	0.995
UK	0.000***	0.348	0.000***	0.563
USA	0.000***	0.392	0.000***	0.227
World ex EA19	0.000***	0.083*	0.000***	0.283
World ex EU27	0.000***	0.474	0.000***	0.580
World	0.000***	0.464	0.000***	0.437

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (14) 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 12: 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	0.058	0.594*	0.600***	0.058	0.495	0.489*
Australia	0.026	0.699**	0.705***	0.020	0.524**	0.374***
Austria	0.014	0.358*	0.330**	0.024	0.611	0.579
Belgium	0.030	0.483**	0.428***	0.028	0.551**	0.480***
Brazil	0.035	0.442***	0.424***	0.048	0.642***	0.593***
Bulgaria	0.027	0.593***	0.207	0.026	0.795**	0.599**
Canada	0.032	0.528*	0.448**	0.041	0.549**	0.520***
China	0.039	0.755*	0.534***	0.038	0.666**	0.494***
Croatia	0.038	0.471***	0.248**	0.031	0.586**	0.476***
Cyprus	0.179	0.581	0.510***	0.262	0.723	0.656*
Czech Rep.	0.034	0.541*	0.409**	0.038	0.586	0.550
Denmark	0.020	0.550***	0.312***	0.017	0.577***	0.559***
Estonia	0.026	0.370***	0.336***	0.027	0.448***	0.405***
Finland	0.040	0.466***	0.546***	0.047	0.487**	0.450***
France	0.035	0.634	0.440***	0.043	0.533	0.547***
Germany	0.026	0.550*	0.424***	0.036	0.514	0.459***
Greece	0.031	0.655***	0.564***	0.028	0.557***	0.479***
Hong Kong, China	0.036	0.420*	0.434**	0.028	0.460*	0.386*
Hungary	0.042	0.596*	0.496*	0.037	0.527	0.512
Iceland	0.087	0.587	0.605***	0.061	0.597**	0.544**
India	0.044	0.952*	0.539***	0.047	0.920	0.690***
Indonesia	0.029	0.600***	0.626***	0.027	0.410***	0.397***
Ireland	0.056	0.511***	0.320**	0.060	0.552**	0.469***
Italy	0.045	0.739*	0.359***	0.042	0.946	0.680**
Japan	0.020	0.443**	0.385***	0.027	0.656**	0.586***
Korea, Republic of	0.026	0.517**	0.580***	0.032	0.511**	0.497**
Latvia	0.018	0.647**	0.299	0.017	0.628**	0.561**
Lithuania	0.027	0.491*	0.465***	0.030	0.744*	0.774*
Luxembourg	0.046	0.567**	0.541***	0.058	0.719	0.695
Malta	0.110	0.807**	0.755*	0.076	0.509***	0.448***
Mexico	0.057	0.301	0.263**	0.065	0.393	0.282
Netherlands	0.017	0.621**	0.494***	0.018	0.568**	0.570***
New Zealand	0.022	0.560**	0.473***	0.026	0.477*	0.413**
N. Macedonia	0.079	0.436	0.438**	0.097	0.528	0.493
Norway	0.034	0.418**	0.411***	0.034	0.451**	0.383*
Poland	0.029	0.573	0.311*	0.027	0.537	0.522***
Portugal	0.037	0.506**	0.523***	0.046	0.630	0.729**
Romania	0.047	0.548*	0.693	0.058	0.670	0.687
Russia	0.024	0.534	0.405**	0.027	0.609	0.455*

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 (21) 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 13: 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	0.017	0.575**	0.465**	0.018	0.640***	0.622***
Adv. econ.	0.023	0.673*	0.485*	0.030	0.859	0.837
CIS	0.021	0.539	0.514	0.025	0.611	0.596
EA 19 extra	0.029	0.627	0.453*	0.029	0.610	0.595
EA 19 intra	0.031	0.630	0.458	0.026	0.521	0.498
EU 27 extra	0.028	0.638	0.460*	0.029	0.648	0.628
EU 27 intra	0.027	0.532	0.388	0.026	0.523	0.500
EU candid.	0.044	0.433***	0.377**	0.056	0.534**	0.504**
EU27	0.027	0.574*	0.846	0.026	0.553	0.476
EURO	0.030	0.551*	0.523***	0.033	0.595	0.580***
EM A. exCNa	0.015	0.440***	0.422**	0.021	0.580***	0.562***
EM exCN	0.014	0.585**	0.666*	0.013	0.558*	0.520*
EM Asia	0.029	0.628*	0.544**	0.023	0.495**	0.487**
EM	0.016	0.538**	0.523**	0.016	0.513**	0.477**
G20 countr.	0.021	0.660**	0.483*	0.023	0.730	0.708
Latin Am.	0.021	0.419*	0.467**	0.025	0.474	0.431*
M.-E./ Afr.	0.015	0.532	0.595	0.017	0.596	0.522*
M.-E./ N-Afr.	0.016	0.423	0.382	0.022	0.549	0.482
Other Asia	0.018	0.520**	0.509***	0.022	0.629***	0.662***
Other CIS	0.045	0.677***	0.598***	0.044	0.656***	0.626**
O. Latin Am.a	0.015	0.486***	0.448***	0.018	0.564***	0.576***
Saudi Arab.	0.030	0.589**	0.360*	0.037	0.673	0.439*
Singapore	0.023	0.618**	0.453**	0.020	0.589***	0.518***
Slovak Rep.	0.053	0.563**	0.449**	0.060	0.627	0.597
Slovenia	0.037	0.656**	0.449***	0.037	0.551**	0.532***
South Afr.	0.077	1.091	0.344***	0.082	1.110	0.392**
Spain	0.042	0.485*	0.276*	0.045	0.407	0.451**
Sub-S. Afr.	0.063	0.471	0.371	0.064	0.468	0.438
Sweden	0.024	0.461**	0.242*	0.022	0.482***	0.398**
Switzerland	0.018	0.569	0.486*	0.026	0.806*	0.777*
Taiwan	0.016	0.830	0.644*	0.017	0.845**	0.785**
Turkey	0.043	0.446***	0.395**	0.055	0.542**	0.476**
UK	0.041	0.545***	0.534***	0.039	0.450***	0.442***
USA	0.034	0.389**	0.275***	0.041	0.467**	0.396**
World ex EA19	0.015	0.576*	0.458**	0.016	0.598**	0.574**
World ex EU27	0.015	0.578**	0.520**	0.014	0.533***	0.503***
World	0.018	0.579*	0.333**	0.017	0.512**	0.488*

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 (21) 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 14: 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***	0.802	0.000***	0.535
Australia	0.000***	0.464	0.000***	0.236
Austria	0.000***	0.800	0.000***	0.360
Belgium	0.000***	0.091*	0.000***	0.928
Brazil	0.000***	0.061*	0.000***	0.865
Bulgaria	0.000***	0.646	0.002***	0.262
Canada	0.000***	0.397	0.000***	0.061*
China	0.000***	0.577	0.000***	0.439
Croatia	0.000***	0.039**	0.000***	0.260
Cyprus	0.000***	0.528	0.000***	0.880
Czech Rep.	0.000***	0.532	0.000***	0.428
Denmark	0.000***	0.247	0.000***	0.594
Estonia	0.000***	0.210	0.000***	0.870
Finland	0.000***	0.059*	0.000***	0.337
France	0.000***	0.293	0.000***	0.080*
Germany	0.000***	0.429	0.000***	0.062*
Greece	0.000***	0.696	0.000***	0.054*
Hong Kong, China	0.000***	0.108	0.000***	0.340
Hungary	0.000***	0.634	0.000***	0.253
Iceland	0.000***	0.994	0.000***	0.263
India	0.000***	0.877	0.001***	0.339
Indonesia	0.000***	0.080*	0.000***	0.590
Ireland	0.000***	0.145	0.000***	0.728
Italy	0.000***	0.794	0.001***	0.028**
Japan	0.000***	0.697	0.000***	0.189
Korea, Republic of	0.000***	0.270	0.000***	0.825
Latvia	0.000***	0.484	0.000***	0.492
Lithuania	0.000***	0.923	0.000***	0.058*
Luxembourg	0.000***	0.058*	0.000***	0.959
Malta	0.000***	0.883	0.000***	0.470
Mexico	0.000***	0.652	0.000***	0.789
Netherlands	0.000***	0.157	0.000***	0.144
New Zealand	0.000***	0.145	0.000***	0.773
N. Macedonia	0.000***	0.378	0.000***	0.000***
Norway	0.000***	0.734	0.000***	0.246
Poland	0.000***	0.055*	0.000***	0.019**
Portugal	0.000***	0.808	0.000***	0.114
Romania	0.000***	0.151	0.000***	0.288
Russia	0.000***	0.423	0.000***	0.215

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (14) 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 15: 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***	0.527	0.000***	0.388
Advanced econ.	0.000***	0.088*	0.000***	0.247
CIS	0.000***	0.712	0.000***	0.126
EA 19 extra	0.000***	0.110	0.000***	0.184
EA 19 intra	0.000***	0.357	0.000***	0.544
EU 27 extra	0.000***	0.033**	0.000***	0.255
EU 27 intra	0.000***	0.278	0.000***	0.928
EU candidates	0.000***	0.457	0.000***	0.019**
EU27	0.000***	0.000***	0.000***	0.340
EURO	0.000***	0.478	0.000***	0.153
Em. Asia ex CN	0.000***	0.299	0.000***	0.259
Em. econ. ex CN	0.000***	0.005***	0.000***	0.194
Emerging Asia	0.000***	0.540	0.000***	0.285
Emerging econ.	0.000***	0.774	0.000***	0.258
G20 countries	0.000***	0.100	0.000***	0.097*
Latin America	0.000***	0.069*	0.000***	0.528
Middle-E./ Afr.	0.000***	0.929	0.000***	0.791
Middle-E./ N- Afr.	0.000***	0.069*	0.000***	0.830
Other Asia	0.000***	0.308	0.000***	0.034**
Other CIS	0.000***	0.139	0.000***	0.357
Other Latin Am.	0.000***	0.487	0.000***	0.989
Saudi Arab.	0.000***	0.096*	0.000***	0.278
Singapore	0.000***	0.039**	0.000***	0.195
Slovak Republic	0.000***	0.103	0.000***	0.357
Slovenia	0.000***	0.567	0.000***	0.949
South Africa	0.002***	0.161	0.031**	0.611
Spain	0.000***	0.000***	0.000***	0.104
Sub-Sah. Africa	0.000***	0.007***	0.000***	0.107
Sweden	0.000***	0.987	0.000***	0.848
Switzerland	0.000***	0.253	0.000***	0.406
Taiwan	0.000***	0.477	0.000***	0.362
Turkey	0.000***	0.782	0.000***	0.032**
UK	0.000***	0.380	0.000***	0.640
USA	0.000***	0.986	0.000***	0.184
World ex EA19	0.000***	0.824	0.000***	0.162
World ex EU27	0.000***	0.746	0.000***	0.015**
World	0.000***	0.498	0.000***	0.765

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (14) 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 Robustness checks

Table 16: Forecast performance of PLS model for unilateral, real imports; 15 day cutoff

	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	0.052	0.693***	0.568***	0.042	0.570***	0.542***
China	0.029	0.629***	0.595***	0.028	0.692**	0.544***
EU27	0.024	0.677**	0.827	0.025	0.684	0.615
France	0.029	0.660*	0.352**	0.040	0.875	0.841
Germany	0.022	0.641**	0.450***	0.018	0.454**	0.353***
Japan	0.021	0.483**	0.316***	0.020	0.687	0.674***
Russia	0.015	0.619**	0.536**	0.015	0.648	0.590*
Sub-Sah. Africa	0.034	0.681***	0.605*	0.032	0.629**	0.530
UK	0.044	0.510**	0.415***	0.048	0.566**	0.516**
USA	0.023	0.714**	0.569***	0.025	0.850	0.756*
World	0.014	0.590*	0.420***	0.014	0.540**	0.464**

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 (21) 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 17: Encompassing tests of PLS and benchmark models for unilateral, real imports; 15 day cutoff

	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***	0.300	0.000***	0.180
China	0.000***	0.052*	0.000***	0.777
EU27	0.000***	0.000***	0.000***	0.077*
France	0.000***	0.753	0.010**	0.000***
Germany	0.000***	0.437	0.000***	0.546
Japan	0.000***	0.146	0.000***	0.083*
Russia	0.000***	0.076*	0.000***	0.318
Sub-Sah. Africa	0.000***	0.860	0.000***	0.343
UK	0.000***	0.404	0.000***	0.045**
USA	0.000***	0.425	0.001***	0.558
World	0.000***	0.753	0.000***	0.731

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (14) 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 18: Forecast performance of PLS model for unilateral, real exports; 15 day cutoff

	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	0.037	0.537***	0.515***	0.043	0.587***	0.543***
China	0.036	0.649*	0.459***	0.047	0.699**	0.519***
EU27	0.041	0.883	1.302	0.038	0.814	0.702
France	0.050	0.892	0.620**	0.054	1.034	1.061
Germany	0.037	0.725	0.559***	0.036	0.703	0.627***
Japan	0.029	0.671**	0.584***	0.031	0.712***	0.637***
Russia	0.030	0.697*	0.529*	0.032	0.764	0.571*
Sub-Sah. Africa	0.069	0.517	0.407	0.074	0.538	0.505
UK	0.047	0.601**	0.588***	0.047	0.629***	0.617***
USA	0.037	0.447**	0.316***	0.033	0.502	0.425*
World	0.017	0.574*	0.330**	0.020	0.779	0.742*

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 (21) 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 19: Encompassing tests of PLS and benchmark models for unilateral, real exports; 15 day cutoff

	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***	0.675	0.000***	0.586
China	0.000***	0.534	0.000***	0.583
EU27	0.000***	0.000***	0.000***	0.295
France	0.011**	0.614	0.140	0.002***
Germany	0.000***	0.490	0.000***	0.007***
Japan	0.000***	0.702	0.000***	0.049**
Russia	0.000***	0.875	0.001***	0.462
Sub-Sah. Africa	0.000***	0.081*	0.000***	0.361
UK	0.000***	0.601	0.000***	0.632
USA	0.000***	0.710	0.000***	0.187
World	0.000***	0.398	0.001***	0.310

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (14) 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 20: 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	0.048	0.786**	0.675***	0.047	0.705***	0.738**
China	0.030	0.658**	0.365***	0.028	0.648**	0.548***
EU27	0.019	0.439**	0.579**	0.027	0.581	0.509
France	0.031	0.403**	0.533***	0.042	0.582	0.482
Germany	0.019	0.437**	0.408***	0.020	0.454**	0.362**
Japan	0.026	0.697**	0.528***	0.024	0.643***	0.589***
Russia	0.019	0.505***	0.564***	0.021	0.501**	0.451***
UK	0.033	0.450***	0.420***	0.041	0.558***	0.546***
USA	0.020	0.622**	0.794	0.019	0.594**	0.542**

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 (21) 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 21: 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***	0.058*	0.000***	0.093*
China	0.000***	0.493	0.000***	0.637
EU27	0.000***	0.007***	0.000***	0.902
France	0.000***	0.000***	0.000***	0.026**
Germany	0.000***	0.807	0.000***	0.845
Japan	0.000***	0.160	0.000***	0.530
Russia	0.000***	0.692	0.000***	0.079*
UK	0.000***	0.361	0.000***	0.207
USA	0.000***	0.000***	0.000***	0.037**

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (14) 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 22: 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	0.038	0.621***	0.596***	0.035	0.634***	0.623***
China	0.055	0.814	0.375***	0.038	0.870	0.412**
EU27	0.027	0.534	0.917	0.035	0.666	0.579
France	0.038	0.498**	0.574***	0.044	0.612*	0.585**
Germany	0.030	0.517	0.598**	0.034	0.612	0.438*
Japan	0.026	0.672**	0.475**	0.026	0.619***	0.644***
Russia	0.036	0.486**	0.391***	0.050	0.774*	0.697**
UK	0.043	0.418***	0.409***	0.050	0.479*	0.459***
USA	0.039	0.708*	1.189	0.041	0.723	0.632

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 (21) 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 23: 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	0.000***	0.250	0.000***	0.322
China	0.000***	0.080*	0.002***	0.848
EU27	0.000***	0.000***	0.000***	0.184
France	0.000***	0.000***	0.000***	0.001***
Germany	0.000***	0.015**	0.000***	0.101
Japan	0.000***	0.002***	0.000***	0.008***
Russia	0.000***	0.201	0.001***	0.920
UK	0.000***	0.473	0.000***	0.941
USA	0.000***	0.000***	0.000***	0.189

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (14) 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 24: Forecast performance of PLS model for unilateral, real imports; two lags

	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	0.048	0.670***	0.357**	0.046	0.588***	0.511***
China	0.024	0.508***	0.500***	0.026	0.615***	0.482***
EU27	0.021	0.582**	0.643*	0.020	0.563	0.510
France	0.034	0.790*	0.338**	0.032	0.878	0.788***
Germany	0.018	0.497**	0.361***	0.022	0.505**	0.410***
Japan	0.020	0.560**	0.409***	0.020	0.576*	0.532***
Russia	0.012	0.411*	0.332**	0.010	0.377**	0.376**
Sub-Sah. Africa	0.035	0.674**	0.601*	0.031	0.544**	0.458*
UK	0.041	0.597**	0.296**	0.040	0.493**	0.474**
USA	0.013	0.369**	0.241**	0.014	0.427**	0.369**
World	0.012	0.394*	0.262***	0.015	0.641	0.535*

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 (21) of the PLS model divided by the standard deviation and the RMSFE of the ARX Benchmark model, respectively. The ARX benchmark model includes leading indicators of the current and two past periods. 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 25: Encompassing tests of PLS and benchmark models for unilateral, real imports; two lags

	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***	0.237	0.000***	0.296
China	0.000***	0.812	0.000***	0.130
EU27	0.000***	0.000***	0.000***	0.039**
France	0.000***	0.064*	0.002***	0.000***
Germany	0.000***	0.362	0.000***	0.438
Japan	0.000***	0.028**	0.000***	0.120
Russia	0.000***	0.152	0.000***	0.792
Sub-Sah. Africa	0.000***	0.085*	0.000***	0.855
UK	0.000***	0.097*	0.000***	0.380
USA	0.000***	0.560	0.000***	0.370
World	0.000***	0.312	0.000***	0.772

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (14) and previous for definitions. The ARX benchmark model includes leading indicators of the current and two past periods. 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 26: Forecast performance of PLS model for unilateral, real exports, two lags

	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	0.035	0.397***	0.347***	0.048	0.639***	0.535***
China	0.039	0.715*	0.505***	0.036	0.585**	0.431***
EU27	0.026	0.573	0.747*	0.027	0.589	0.515
France	0.034	0.628	0.392***	0.042	0.533	0.548***
Germany	0.026	0.550*	0.561***	0.035	0.516	0.483***
Japan	0.021	0.493**	0.357***	0.026	0.656**	0.560***
Russia	0.027	0.615	0.316	0.026	0.609	0.433*
Sub-Sah. Africa	0.062	0.468	0.377	0.065	0.480	0.443
UK	0.040	0.541***	0.478***	0.038	0.450**	0.435***
USA	0.034	0.388**	0.333***	0.040	0.457**	0.361**
World	0.018	0.561*	0.268**	0.018	0.483**	0.470*

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 (21) of the PLS model divided by the standard deviation and the RMSFE of the ARX Benchmark model, respectively. The ARX benchmark model includes leading indicators of the current and two past periods. 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 27: Encompassing tests of PLS and benchmark models for unilateral, real exports, two lags

	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***	0.000***	0.000***	0.313
China	0.000***	0.523	0.000***	0.689
EU27	0.000***	0.000***	0.000***	0.192
France	0.000***	0.133	0.000***	0.020**
Germany	0.000***	0.479	0.000***	0.039**
Japan	0.000***	0.354	0.000***	0.034**
Russia	0.000***	0.758	0.000***	0.350
Sub-Sah. Africa	0.000***	0.006***	0.000***	0.439
UK	0.000***	0.530	0.000***	0.755
USA	0.000***	0.352	0.000***	0.354
World	0.000***	0.217	0.000***	0.735

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. The ARX benchmark model includes leading indicators of the current and two past periods. See equations (14) 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 28: Forecast performance of PLS model for unilateral, real imports

	h=0			h=1		
	Abs. RMSFE PLS	rel. RMSFE Std. Dev.	rel. RMSFE Benchmark	Abs. RMSFE PLS	rel. RMSFE Std. Dev.	rel. RMSFE Benchmark
Brazil	0.048	0.670***	0.661***	0.046	0.588***	0.623***
China	0.024	0.508***	0.464***	0.027	0.615***	0.506***
EU27	0.021	0.587**	0.681**	0.020	0.563	0.556
France	0.031	0.608*	0.589*	0.037	0.771	0.716
Germany	0.019	0.511**	0.430***	0.021	0.497**	0.391**
Japan	0.020	0.540**	0.527**	0.021	0.581*	0.585*
Russia	0.012	0.411*	0.404*	0.011	0.435**	0.401*
Sub-Sah. Africa	0.035	0.679*	0.657*	0.029	0.553**	0.460
UK	0.039	0.544**	0.496**	0.037	0.466***	0.436***
USA	0.014	0.369**	0.319***	0.014	0.427**	0.391**
World	0.012	0.394*	0.393**	0.015	0.637	0.682

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 (21) of the PLS model divided by indicator 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 indicator-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 29: Encompassing tests of PLS and benchmark models for unilateral, real imports; leading indicators only

	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***	0.299	0.000***	0.155
China	0.000***	0.802	0.000***	0.163
EU27	0.000***	0.000***	0.000***	0.206
France	0.000***	0.527	0.000***	0.001***
Germany	0.000***	0.327	0.000***	0.681
Japan	0.000***	0.091*	0.000***	0.060*
Russia	0.000***	0.317	0.000***	0.561
Sub-Sah. Africa	0.000***	0.073*	0.000***	0.696
UK	0.000***	0.686	0.000***	0.450
USA	0.000***	0.961	0.000***	0.309
World	0.000***	0.848	0.000***	0.796

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (14) 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 30: Forecast performance of PLS model for unilateral, real exports; leading indicators only

	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	0.035	0.442***	0.413***	0.048	0.642***	0.662***
China	0.039	0.755*	0.614**	0.038	0.666**	0.581**
EU27	0.027	0.574*	0.660*	0.026	0.553	0.537
France	0.035	0.634	0.607	0.043	0.533	0.576
Germany	0.026	0.550*	0.582*	0.036	0.514	0.464*
Japan	0.020	0.443**	0.497***	0.027	0.656**	0.632**
Russia	0.024	0.534	0.458	0.027	0.609	0.467
Sub-Sah. Africa	0.063	0.471	0.368	0.064	0.468	0.453
UK	0.041	0.545***	0.516***	0.039	0.450***	0.431***
USA	0.034	0.389**	0.359***	0.041	0.467**	0.471
World	0.018	0.579*	0.447**	0.017	0.512**	0.542**

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 (21) of the PLS model divided by indicator 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 indicator-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 31: Encompassing tests of PLS and benchmark models for unilateral, real exports; leading indicators only

	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***	0.016**	0.000***	0.236
China	0.000***	0.809	0.000***	0.857
EU27	0.000***	0.036**	0.000***	0.870
France	0.000***	0.342	0.000***	0.304
Germany	0.000***	0.957	0.000***	0.183
Japan	0.000***	0.072*	0.000***	0.903
Russia	0.000***	0.469	0.000***	0.236
Sub-Sah. Africa	0.000***	0.288	0.000***	0.988
UK	0.000***	0.946	0.000***	0.612
USA	0.000***	0.932	0.000***	0.085*
World	0.000***	0.957	0.000***	0.904

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (14) 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 32: Forecast performance of PLS model for unilateral, real imports; current period indicators

	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	0.045	0.632***	0.573***	0.046	0.586***	0.580***
China	0.024	0.503***	0.391***	0.027	0.609***	0.476***
EU27	0.021	0.593**	0.910	0.020	0.563	0.497
France	0.030	0.401*	0.496***	0.037	0.493	0.316
Germany	0.018	0.497**	0.375***	0.022	0.493**	0.310***
Japan	0.020	0.509**	0.308**	0.021	0.517*	0.445***
Russia	0.012	0.524*	0.475**	0.011	0.436**	0.409**
Sub-Sah. Africa	0.035	0.679*	0.541**	0.029	0.546**	0.496*
UK	0.041	0.626**	0.598**	0.039	0.475**	0.414**
USA	0.014	0.374**	0.387***	0.014	0.447**	0.406***
World	0.012	0.403*	0.452***	0.015	0.641	0.511

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 (21) 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 33: Encompassing tests of PLS and benchmark models for unilateral, real imports; current period indicators

	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***	0.674	0.000***	0.713
China	0.000***	0.807	0.000***	0.093*
EU27	0.000***	0.000***	0.000***	0.094*
France	0.000***	0.904	0.000***	0.088*
Germany	0.000***	0.977	0.000***	0.228
Japan	0.000***	0.126	0.000***	0.616
Russia	0.000***	0.058*	0.000***	0.946
Sub-Sah. Africa	0.000***	0.558	0.000***	0.913
UK	0.000***	0.010**	0.000***	0.131
USA	0.000***	0.201	0.000***	0.616
World	0.000***	0.763	0.000***	0.281

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (14) 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 34: Forecast performance of PLS model for unilateral, real exports, no lag adjustment for leading indicators; current period indicators

	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	0.035	0.438***	0.507***	0.046	0.582***	0.498***
China	0.039	0.726*	0.512***	0.039	0.729**	0.553***
EU27	0.027	0.573	1.200	0.027	0.582	0.497
France	0.037	0.505	0.773**	0.043	0.436	0.264
Germany	0.026	0.520*	0.732***	0.035	0.457	0.235*
Japan	0.021	0.525**	0.541***	0.027	0.667**	0.599**
Russia	0.024	0.538	0.388*	0.027	0.602	0.446*
Sub-Sah. Africa	0.062	0.468	0.505	0.066	0.480	0.406
UK	0.041	0.545***	0.533***	0.039	0.451***	0.448***
USA	0.034	0.429**	0.656	0.040	0.740**	0.560**
World	0.018	0.558*	0.484**	0.018	0.760**	0.562

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 (21) 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 35: Encompassing tests of PLS and benchmark models for unilateral, real exports; current period indicators

	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***	0.603	0.000***	0.047**
China	0.000***	0.298	0.000***	0.821
EU27	0.000***	0.000***	0.000***	0.302
France	0.000***	0.229	0.000***	0.346
Germany	0.000***	0.394	0.000***	0.312
Japan	0.000***	0.010**	0.000***	0.006***
Russia	0.000***	0.566	0.000***	0.150
Sub-Sah. Africa	0.000***	0.201	0.000***	0.570
UK	0.000***	0.339	0.000***	0.886
USA	0.000***	0.204	0.000***	0.000***
World	0.000***	0.853	0.000***	0.002***

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (14) 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 36: Forecast performance of PLS model for unilateral, real imports, longer tie period

	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	0.045	0.595***	0.537***	0.046	0.599***	0.610***
China	0.025	0.519***	0.514***	0.027	0.567***	0.540***
EU27	0.022	0.622**	0.760	0.020	0.563	0.506
France	0.034	0.555*	0.484**	0.032	0.518	0.485
Germany	0.018	0.560**	0.641***	0.022	0.660**	0.539*
Japan	0.020	0.461**	0.436**	0.021	0.492*	0.490*
Russia	0.012	0.573*	0.489*	0.011	0.479**	0.415**
Sub-Saharan Africa	0.035	0.679*	0.603*	0.029	0.555**	0.467*
UK	0.041	0.579**	0.515**	0.040	0.553**	0.528***
USA	0.014	0.472**	0.422***	0.014	0.472**	0.461**
World	0.012	0.580*	0.429*	0.015	0.721	0.717

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 (21) 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 37: Encompassing tests of PLS and benchmark models for unilateral, real imports, longer training period

	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***	0.300	0.000***	0.252
China	0.000***	0.136	0.000***	0.431
EU27	0.000***	0.000***	0.000***	0.086*
France	0.000***	0.873	0.000***	0.015**
Germany	0.000***	0.051*	0.000***	0.652
Japan	0.000***	0.935	0.000***	0.119
Russia	0.000***	0.001***	0.000***	0.175
Sub-Saharan Africa	0.000***	0.478	0.000***	0.844
UK	0.000***	0.800	0.000***	0.400
USA	0.000***	0.482	0.000***	0.539
World	0.000***	0.181	0.000***	0.094*

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (14) 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 38: Forecast performance of PLS model for unilateral, real exports, two lags, longer time period

	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	0.035	0.504***	0.498***	0.048	0.660***	0.688***
China	0.039	0.537*	0.512**	0.037	0.510**	0.505**
EU27	0.027	0.573	0.845	0.028	0.589	0.507
France	0.037	0.573	0.494*	0.043	0.659	0.620
Germany	0.026	0.521*	0.614*	0.035	0.683	0.536
Japan	0.020	0.527**	0.470***	0.027	0.674**	0.658**
Russia	0.025	0.602	0.455	0.027	0.637	0.593
Sub-Sah. Africa	0.059	0.441	0.347	0.066	0.480	0.450
UK	0.041	0.594***	0.521***	0.039	0.548**	0.533***
USA	0.034	0.715**	0.678***	0.041	0.832**	0.844***
World	0.018	0.676*	0.510*	0.017	0.603**	0.597**

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 (21) 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 39: Encompassing tests of PLS and benchmark models for unilateral, real exports, longer time period

	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***	0.000***	0.000***	0.313
China	0.000***	0.523	0.000***	0.689
EU27	0.000***	0.000***	0.000***	0.192
France	0.000***	0.133	0.000***	0.020**
Germany	0.000***	0.479	0.000***	0.039**
Japan	0.000***	0.354	0.000***	0.034**
Russia	0.000***	0.758	0.000***	0.350
Sub-Sah. Africa	0.000***	0.006***	0.000***	0.439
UK	0.000***	0.530	0.000***	0.755
USA	0.000***	0.352	0.000***	0.354
World	0.000***	0.217	0.000***	0.735

Significance of the Wald test implies rejection of the null hypothesis and indicates not encompassed predictive power of the underlying model. See equations (14) 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 40: Comparison of various machine learning tools with and without sparcity against the sparse PLS; real exports

	Sparse PLS abs.RMSFE	Lasso rRMSFE	PCR rRMSFE	Reg. Tree rRMSFE	PLS rRMSFE	Sparse Lasso rRMSFE	Sparse PCR rRMSFE	Sparse Tree rRMSFE
h=0								
Brazil	0.037	0.524***	0.517***	0.456***	0.495***	0.640***	0.895*	0.514***
China	0.039	0.521*	0.481**	0.480**	0.501**	0.567*	0.876**	0.489*
EU27	0.027	0.557*	0.538*	0.544*	0.517	0.572	0.652	0.524**
France	0.038	0.588	0.552*	0.532*	0.535	0.602	0.779	0.576**
Germany	0.026	0.530	0.453*	0.501*	0.472	0.564	1.071	0.521**
Japan	0.021	0.525**	0.520***	0.500**	0.458***	0.525**	0.900**	0.516***
Russia	0.028	0.662	0.664	0.666	0.682	0.662	0.776	0.685
Sub-Sah. Africa	0.061	0.429*	0.467	0.416	0.512*	0.500**	0.992	0.437
UK	0.040	0.579***	0.574***	0.532***	0.554***	0.598***	0.968	0.592***
USA	0.034	0.690**	0.646**	0.667**	0.625**	0.687**	0.952	0.669**
World	0.018	0.669*	0.624**	0.624**	0.629**	0.671*	0.967	0.648***
h=1								
Brazil	0.048	0.654***	0.665***	0.627***	0.661***	0.713***	0.978	0.591***
China	0.038	0.537**	0.518**	0.450***	0.525**	0.559**	0.992	0.483***
EU27	0.027	0.590	0.557	0.530	0.524	0.572	0.679	0.572
France	0.042	0.653	0.594	0.642	0.594	0.653	0.869	0.642
Germany	0.036	0.689	0.620	0.622	0.616	0.701	0.973	0.624
Japan	0.027	0.649**	0.562	0.679**	0.571	0.672**	0.909	0.703**
Russia	0.027	0.649	0.654	0.664	0.674	0.651	0.842	0.648
Sub-Sah. Africa	0.064	0.450	0.448	0.412	0.520	0.464	0.846	0.464
UK	0.039	0.560**	0.564**	0.516***	0.561**	0.593**	0.963	0.556***
USA	0.037	0.765*	0.747*	0.724	0.663	0.764*	0.845	0.684***
World	0.017	0.610**	0.577*	0.628*	0.535*	0.607**	1.080	0.658**

Note: The column "Sparse PLS, abs. RMSFE" measures the absolute forecast error of the sparse partial least squares model. Columns labeled "rRMSFE" measure the relative RMSFE of the sparse PLS as compared to the noted machine learning tools. "Sparse" always refers to prior shrinkage of the predictor space using LASSO.

A.3 Additional material

Table 41: Variables for estimation of US-American imports in current month

Course/Orientation	Description
Eastern course	Western Panama Canal
Southern course	Western Mexican Coast
Northern course	Gulf of Mexico
Northern course	Central East Coast USA
Northern course	Northern East Coast USA
Western course	Strait of Gibraltar
Southern course	West Coast of Great Britain
Northern course	North Sea
Northern course	Java Sea
Southern course	Southern Chinese Coast
Arrival	Miami, United States
Arrival	New York, United States
Arrival	Freeport, Bahamas
Arrival	Cartagena, Colombia
Arrival	Hamburg, Germany
Arrival	Kiel, Germany
Arrival	Dunkirk, France
Arrival	Trieste, Italy
Arrival	Ain Sukhna, Egypt
Arrival	Chennai, India
Arrival	Tokyo, Japan
Arrival	Kobe, Japan
Departure	Jackson, United States
Departure	All ports, United States
Departure	Sepetiba, Brazil
Departure	Colon, Panama
Departure	Halifax, Canada
Departure	Bremen, Germany
Departure	Hamburg, Germany
Departure	La Spezia, Italy
Departure	Larvik, Norway
Departure	Chennai, India
Departure	Dong Guan, China

The order of variables does no reflect the variable importance in the final estimation of the PLS.

Table 42: Overview of leading indicators used in benchmark model

Country	Leading indicators
Argentina	Argentina metr. t. - IMF, US IP, GSPCI - NY Fed
Australia	Australia metr. t. - IMF, US IP, GSPCI - NY Fed
Austria	Austria COF, Austria ESI, US IP, GSPCI - NY Fed
Belgium	Belg. COF, Belg. ESI, Belg. metr. t. - IMF, US IP, GSPCI - NY Fed, Northrange RWI/ISL
Brazil	Brazil metr. t. - IMF, US IP, GSPCI - NY Fed
Bulgaria	Bulgaria COF, Bulgaria ESI, Bulgaria metr. t. - IMF, US IP, GSPCI - NY Fed
Canada	Canada metr. t. - IMF, US IP, GSPCI - NY Fed
China	China metr. t. - IMF, US IP, GSPCI - NY Fed
Croatia	Croatia COF, Croatia ESI, Croatia metr. t. - IMF, US IP, GSPCI - NY Fed
Cyprus	Cyprus COF, Cyprus ESI, Cyprus metr. t. - IMF, US IP, GSPCI - NY Fed
Czech Republic	Czech Republic COF, Czech Republic ESI, US IP, GSPCI - NY Fed
Denmark	Denmark COF, Denmark ESI, Denmark metr. t. - IMF, US IP, GSPCI - NY Fed
EU27	EU27 COF, EU27 ESI, US IP, GSPCI - NY Fed
EURO	EURO COF, EURO ESI, EURO metr. t. - IMF, US IP, GSPCI - NY Fed
Estonia	Estonia COF, Estonia ESI, Estonia metr. t. - IMF, US IP, GSPCI - NY Fed
Finland	Finland COF, Finland ESI, Finland metr. t. - IMF, US IP, GSPCI - NY Fed
France	France COF, France ESI, France metr. t. - IMF, US IP, GSPCI - NY Fed
Germany	DE COF, DE ESI, DE metr. t. - IMF, US IP, GSPCI - NY Fed
Greece	Greece COF, Greece ESI, Greece metr. t. - IMF, US IP, GSPCI - NY Fed
Hong Kong, China	Hong Kong, China metr. t. - IMF, US IP, GSPCI - NY Fed
Hungary	Hungary COF, Hungary ESI, US IP, GSPCI - NY Fed
Iceland	Iceland metr. t. - IMF, US IP, GSPCI - NY Fed
India	India metr. t. - IMF, US IP, GSPCI - NY Fed
Indonesia	Indonesia metr. t. - IMF, US IP, GSPCI - NY Fed
Ireland	Ireland COF, Ireland ESI, Ireland metr. t. - IMF, US IP, GSPCI - NY Fed
Italy	Italy COF, Italy ESI, Italy metr. t. - IMF, US IP, GSPCI - NY Fed
Japan	Japan metr. t. - IMF, US IP, GSPCI - NY Fed
Korea, Republic of	Korea, Republic of metr. t. - IMF, US IP, GSPCI - NY Fed
Latvia	Latvia COF, Latvia ESI, Latvia metr. t. - IMF, US IP, GSPCI - NY Fed
Lithuania	Lithuania COF, Lithuania ESI, Lithuania metr. t. - IMF, US IP, GSPCI - NY Fed
Luxembourg	Luxembourg COF, Luxembourg ESI, US IP, GSPCI - NY Fed
Malta	Malta COF, Malta ESI, Malta metr. t. - IMF, US IP, GSPCI - NY Fed
Mexico	Mexico metr. t. - IMF, US IP, GSPCI - NY Fed
Netherlands	NL COF, NL ESI, NL metr. t. - IMF, US IP, GSPCI - NY Fed, Northrange RWI/ISL
New Zealand	New Zealand metr. t. - IMF, US IP, GSPCI - NY Fed
North Macedonia	North Macedonia COF, North Macedonia ESI, US IP, GSPCI - NY Fed
Norway	Norway metr. t. - IMF, US IP, GSPCI - NY Fed
Poland	Poland COF, Poland ESI, Poland metr. t. - IMF, US IP, GSPCI - NY Fed
Portugal	Portugal COF, Portugal ESI, Portugal metr. t. - IMF, US IP, GSPCI - NY Fed
Romania	Romania COF, Romania ESI, US IP, GSPCI - NY Fed
Russian Federation	Russian Federation metr. t. - IMF, US IP, GSPCI - NY Fed
Saudi_Arabia	Saudi_Arabia metr. t. - IMF, US IP, GSPCI - NY Fed
Singapore	Singapore metr. t. - IMF, US IP, GSPCI - NY Fed
Slovak Republic	Slovak Republic COF, Slovak Republic ESI, US IP, GSPCI - NY Fed
Slovenia	Slovenia COF, Slovenia ESI, Slovenia metr. t. - IMF, US IP, GSPCI - NY Fed
South Africa	South Africa metr. t. - IMF, US IP, GSPCI - NY Fed
Spain	Spain COF, Spain ESI, Spain metr. t. - IMF, US IP, GSPCI - NY Fed
Sweden	Sweden COF, Sweden ESI, Sweden metr. t. - IMF, US IP, GSPCI - NY Fed
Turkey	Turkey COF, Turkey ESI, Turkey metr. t. - IMF, US IP, GSPCI - NY Fed
United Kingdom	UK COF, UK ESI, UK metr. t. - IMF, US IP, GSPCI - NY Fed
United States	United States metr. t. - IMF, US IP, GSPCI - NY Fed
World	World metr. t. - IMF, US IP, GSPCI - NY Fed

Note: metr. t. - IMF, refers to maritime export or import in metric tons as calculated by IMF. COF and ESI refer to the industrial confidence and economic sentiment indicators by the European Commission. US IP refers to the industrial product of the United States. GSCPI refers to the global supply chain pressure index by the New York Federal Reserve. Northrange RWI/ISL refers to the Northrange container throughput indicator by RWI/ISL.