ISSN 1936-5330



Introducing Trimmed Imports and Exports

Thomas R. Cook, Mariia Dzholos, and Johannes Matschke

August 2026
RWP 26-08
https://doi.org/10.18651/RWP2026-08



Introducing Trimmed Imports and Exports*

Thomas Cook
FRB Kansas City

Mariia Dzholos
University of Zurich

Johannes Matschke
FRB Kansas City

August 17, 2026

Abstract

Trade flows are volatile and notoriously difficult to predict. This paper introduces trimmed import and export growth rates for the United States to help improve trade forecasts. We calculate the growth rates as a trimmed mean that strips out subcomponents with less predictive power. By separating the signal from the noise, our trimmed growth rates provide a better perspective on where aggregate imports or exports are heading, particularly at horizons four to 12 months ahead. Consequently, out-of-sample forecasts based on our trimmed growth rates reduce forecast errors by up to 15 percent relative to forecasts from aggregate import or export growth. These trimmed forecasts are simple to compute, outperform both naive forecasts and models that rely on a richer information set, and are particularly useful during episodes with heightened volatility. We update the trimmed growth rates monthly and make them publicly available.

JEL Classification: **JEL: F17; C22**

Keywords: Imports and Exports; Forecasting; Signal Extraction

*We thank the Census Bureau and our colleagues at the Kansas City Fed for helpful discussions. The views expressed are those of the authors and do not necessarily reflect the positions of the Federal Reserve Bank of Kansas City or the Federal Reserve System.

1. Introduction

A defining feature of imports and exports is their volatility. For example, in anticipation of new U.S. tariffs, firms increased inventories leading to an import surge of 38 percent annualized in the first quarter of 2025 relative to the previous quarter. However, even outside major trade policy announcements, or similarly turning points in business cycles, growth rates for imports and exports frequently change by more than 20 percentage points annualized from one quarter to the next.[1] These large swings make it difficult to predict imports and exports as evident from published trade forecasts and even sophisticated models tend to have large forecast errors.[2]

In this article, we introduce monthly nominal trimmed import and export growth rates for the United States. The trimmed growth rates are calculated from disaggregated import and export data at the 2-digit HS level for goods and include survey data on services. We trim components with lower predictive power and calculate a weighted average over the remaining components. Subsequently, we show that trimmed growth rates improve forecasts for aggregate imports and exports out-of-sample based on simple bivariate regressions ('trimmed forecasts'), particularly over horizons four to 12 months ahead.

The appeal of trimmed imports and trimmed exports relates to two factors: First, the correlation between current trimmed growth and future aggregate growth is higher than the serial correlation of the aggregate growth rate itself. For example at the 12 month horizon, the correlation between current trimmed import growth and future aggregate import growth is about three times as high and the correlation between current trimmed export growth and future aggregate export growth is about six times as high. Trimmed growth is therefore a better indicator as to where the aggregate index

[1] Figure A1 in the appendix shows annualized quarterly nominal growth rates for imports, exports, and consumption. Imports and exports are much more volatile than consumption and – though not shown in the chart – about as volatile as investment. Swings are particularly large around recessions when imports and exports decline sharply, just to rebound afterwards.

[2] Table B1 in the appendix compares import and export forecasts with a consumption forecast, all of which are from GDPNow (Atlanta Fed). The table has two takeaways: First, import and export forecasts for the current quarter are about five times as imprecise than the consumption forecast for the same quarter as measured by the mean squared error (MSE). Second, the one quarter ahead forecasts are about 20-40 times as imprecise as the current quarter consumption forecast.

is heading than the aggregate series itself. Second, the correlation between trimmed growth and subsequent aggregate growth is less noisy throughout time than the serial correlation of the aggregate growth rate. This helps with more precisely estimating the relationship in-sample, while also improving the out-of-sample fit as one reduces forecast errors from extrapolating the average correlation. In line with this feature, we show that trimmed forecasts are quantitatively most useful when imports and exports swing notably. However, trimmed forecasts also tend to outperform other forecasts during 'normal' times.

Regarding forecast accuracy, the trimmed forecasts are best used at horizons four to 12 months ahead during which they improve the out-of-sample fit by up to 15 percent. These gains are statistically significant at the 1% level and measured relative to forecasts based on the same statistical model, but with total import or export growth as the only explanatory variable. On the flip side, the trimmed forecasts are less appealing for the near-term. Indeed, GDPNow provides the best near-term forecast across our comparison group, in line with work by Faust and Wright (2009). This should not surprise considering that GDPNow utilizes a large information set, while mimicking the methods used by the BEA to construct quarterly GDP and its subcomponents.

We compare the trimmed forecasts to several alternatives: they notably outperform simple rule-of-thumb forecasts like a random walk (no change forecast), or moving and historical averages. The trimmed index also outperforms a weighted median, which we calculate from disaggregated import or export data. The weighted median is a well established trend measure outside trade and, for example, used for inflation (Cleveland Fed). We also show that trimmed forecasts are more precise out-of-sample than forecasts generated based on predictive regressions involving several other variables simultaneously like new orders, industrial production, or import and export prices that are commonly used in forecasting models. Last but not least, trimmed forecasts are also more reliable than forecasts from a state-of-the-art Bayesian variable selection model (Koop and Korobilis, 2023). That said, despite notable improvements, forecasting imports and exports remains challenging and point estimates can be far off.

Our approach to strip out subcomponents of an aggregate index is inspired by the Dallas Fed trimmed mean PCE inflation rate. We first calculate monthly seasonally adjusted growth rates for the in total close to 110 subcomponents of the aggregate import and export indices. We then trim components with low and high growth rates. The trimmed growth rate is a re-weighted and normalized average over the remaining subcomponents.

We mostly trim at the bottom of the import and export growth distribution. Specifically, we trim components below the 65th percentile and above the 94th percentile, which maximizes the serial correlation between current trimmed growth and subsequent aggregate growth over the next six to 12 months ahead.[3] Intuitively, drops in aggregate imports or exports tend to be temporary in contrast to more persistent surges (see Figure A1 in the appendix). Disaggregated negative or close to zero import or export growth rates are therefore not indicative as to where overall imports or exports are heading. This feature also explains why a (weighted) median contains less of a signal than the trimmed series for imports and exports.

By trimming mostly from below and by removing 70 percent of the components, the weighted (but not normalized) trimmed growth rate is higher and more volatile than the aggregate growth rate. We adjust the volatility of the trimmed growth rate by multiplying the series with (the square root of) the share of retained components. This adjustment controls for the different number of components. Consequently, the trimmed growth rate is about 35 percent smoother than the aggregate growth rate, precisely because the retained components are on average less volatile than the omitted components. We also adjust the mean of the trimmed growth rate to have the same mean as the aggregate series. These normalizations have no effect on the usefulness of the trimmed growth rate as a predictor, but facilitate a direct comparison with the aggregate growth rate.

Related Literature. We introduce trimmed import and export growth rates to improve trade forecasts. Related papers have evaluated leading indicators from high frequency

[3]For reference, the trimmed mean PCE inflation rate by the Dallas Fed omits categories below the 24th percentile and above the 69th percentile in the price change distribution. Because the distribution is negatively skewed on average, the re-weighted average has about the same mean as headline inflation.

shipping data, surveys, or international trade flows (see, for example, Grimme et al., 2021, Lehmann, 2021, Grossmann and Scheufele, 2019). Other work in contrast focuses on methodological contributions (see, for example, D'Agostino et al., 2017, Golinelli and Parigi, 2014, Cushman, 1990). We instead filter the subcomponents of total imports and exports and estimate a series of OLS regressions. Moreover, we argue that this simple method improves forecasts over forecast horizons up to one year.

Our work also relates to the literature on whether disaggregated data can improve aggregate forecasts. In particular, we show that trimmed imports and exports, which are based on disaggregated data, improve forecasts for total imports and exports. As such, this finding is in line with Hendry and Hubrich (2011), who argue that combining disaggregated data can help forecast an aggregate. Debates over whether to compose forecasts directly on the aggregate series of interest or to produce forecasts aggregated from component series date back to at least the mid 1960s (Dunn et al., 1976). The presumed advantage of forecasting subcomponents and then aggregating is that it can generate more accurate forecasts depending on the volatility and covariance structure (Athanasopoulos et al., 2019). To address this, a variety of aggregation and reconciliation methods have been proposed, but at a practical level, we follow Bryan and Cecchetti (1994) on core inflation. Specifically, we remove certain subcomponents so that the aggregation of the remaining subcomponents is a more stable indicator of the trend in the aggregate series.

The remainder of the paper is structured as follows. Section 2 develops the trimmed import and export growth rates and highlights their performance out-of-sample. Section 3 compares trimmed forecasts with alternative forecasts. Section 4 explains when trimmed forecasts are most useful. Section 5 concludes.

2. Trimmed Imports and Exports

This section presents the main contribution. Section 2.1 introduces trimmed import and export growth rates. Section 2.2 shows out-of-sample gains from using trimmed growth rates. Section 2.3 explains why the trimmed growth rate is a better indicator

for future aggregate growth than the aggregate growth rate itself.

2.1. Construction

Data. We decompose aggregate imports and exports as follows: we use subcomponents at the 2-digit HS level from the U.S. Census Bureau's USA Trade database for goods and complement this data with survey responses on services from the BEA. In total, we end up with 109 categories for imports and 108 categories for exports both of which fully characterize the respective aggregate index.[4] The HS classification describes goods based on their attributes. In terms of granularity, subcomponents include, for example, 'iron and steel', 'clocks and watches', or 'financial services'. All subcomponents are nominal, reported in U.S. dollars and available since January 2002.

Seasonal Adjustment. The raw data for goods at the 2-digit HS level are not seasonally adjusted. The U.S. Census Bureau seasonally adjusts imports and exports at the 5-digit *end-use* level – which classifies shipments based on their use rather than their physical characteristics – and aggregates the individual series to produce seasonally adjusted total imports and exports.[5] The official procedure for seasonal adjustment at the end-use level involves three steps: an algorithm based on X-13ARIMA-SEATS, rare subjective adjustments to the individual series, and forcing. The latter ensures that seasonal effects cancel over the year so that the annual total for the seasonally adjusted series matches the non-seasonally adjusted series. Both end-use and HS codes are conceptually valid classifications for our purpose, however granular end-use data are not publicly available. We therefore resort to the HS classification and follow similar seasonal adjustment procedures. In particular, we run X-13ARIMA-SEATS at the 2-digit HS level, force the seasonally adjusted data to match the non-seasonally adjusted series at annual frequency, and mechanically set any rare negative entry for the seasonally adjusted series to a marginally positive value.[6] Figure A2 in the appendix compares

[4]The difference comes from 'special import classifications', a category which is only present for imports but not exports. Otherwise, the classifications coincide.

[5]To provide an example, suppose a manufacturer imports 'steel gear'. The HS code describes the part itself. The end-use code assigns the product for 'use in manufacturing agricultural equipment'.

[6]It is important to incorporate trading day effects. For example, trading days in February tend to be lower than in other months. This reduces imports and exports irrespective of seasonal effects.

seasonally adjusted monthly nominal growth rates of total imports and exports based on the official BEA release with our own seasonally adjusted aggregate growth rate form the 2-digit HS classification including services. Both series essentially overlap with a correlation of 0.97 for imports and exports.

Methodology. We construct trimmed import and export growth rates as follows: Let $x_t \in \{IM_t, EX_t\}$ denote the total import or export index in year-month t. Subcomponents are referenced by i. Therefore:

$$x_t = \sum_i x_{i,t}. \tag{1}$$

Defining the month-over-month growth rate from $t-1$ to t as $\hat{x}_{t|t-1}$ and the weight of subcomponent i in the overall measure during year-month t as $\omega_{i,t}$, we can express the aggregate growth rate as:

$$\hat{x}_{t|t-1} = \sum_i \omega_{i,t-1}\hat{x}_{i,t|t-1}. \tag{2}$$

We next rank the subcomponents based on their growth rates each month and trim components if their monthly growth rate is above ξ_h or below ξ_l.[7] As discussed below, these parameters will correspond to percentiles of the monthly import or export growth distribution across all subcomponents. To save notation we subsequently reference these thresholds by $\Xi = \{\xi_h, \xi_l\}$. We define the set of all *remaining* components as $\mathbb{C}_\Xi$. With these conventions, the month-over-month growth rate of the trimmed index $\hat{x}^{TM}_{t|t-1}(\Xi)$ equals:

$$\hat{x}^{TM}_{t|t-1}(\Xi) = \sum_{i \in \mathbb{C}_\Xi} \overline{\omega}_{i,t-1}\hat{x}_{i,t|t-1}, \tag{3}$$

with $\overline{\omega}_{i,t} = \frac{\omega_{i,t}}{\sum_{k \in \mathbb{C}_\Xi} \omega_{k,t}}$. The weight adjustment ensures that all remaining subcomponents sum up to 100 percent. Last but not least, we normalize the trimmed growth rate along two dimensions: First, we multiply the trimmed growth rate by the square root

[7]The purpose of this paper is to produce a trimmed growth rate that improves forecasts for the aggregate index. With this in mind, we also checked if it is more desirable to trim based on weighted rather than unweighted growth rates. This was not the case.

of the share of retained subcomponents. This controls for the fewer components in the trimmed series, which would otherwise mechanically increase the volatility. Second, we match the mean of the aggregate growth rate. Both adjustments have no effect on the usefulness of the trimmed growth rate as a predictor, but they facilitate a direct comparison with the aggregate series.

Calibration. We select ξ_h and ξ_l to maximize the predictive power of the trimmed growth rate. In particular, we calibrate the two parameters based on a series of bivariate regressions that resemble Local Projections (Òscar Jordà, 2005):

$$\hat{x}_{t+h|t} = \alpha^h + \beta^h \hat{x}^{TM}_{t|t-1}(\Xi) + \epsilon^{TM}_{t+h}. \tag{4}$$

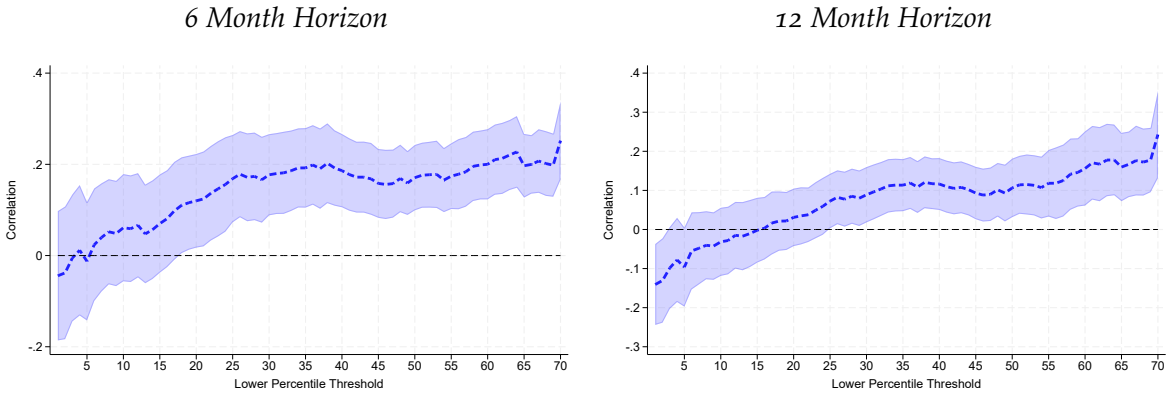
The dependent variable $\hat{x}_{t+h|t}$ is the h month cumulative nominal growth rate of the aggregate import or export index between year-month $t+h$ and t. The independent variable is the monthly nominal growth rate of the trimmed import or export index between t and $t-1$. As described previously, the trimmed growth rate is a function of Ξ and thus the components that are trimmed. We estimate this equation over horizons up to 12 months and for various Ξ. We standardize both the dependent and independent variable: β^h therefore represents the correlation coefficient between the trimmed growth rate in period t and the cumulative aggregate growth rate over the subsequent h months. Since multi-step ahead forecast errors are serially correlated we utilize Newey-West standard errors with lags equal to the forecasting horizon.

When calibrating ξ_h and ξ_l we have two targets in mind. First, we maximize the serial correlation between current monthly trimmed growth and subsequent aggregate growth over the next six to 12 months ahead. Underlying this approach is extensive testing, which suggest that trimmed forecasts are most useful over the medium term. Second, we retain 30 percent of the subcomponents for the trimmed index. Including more components decreases the forecasting performance.[8] Last but not least, we calibrate the cutoffs based on the entire sample, though we obtain the same thresholds if we instead followed out-of-sample procedures that we describe in the next section.

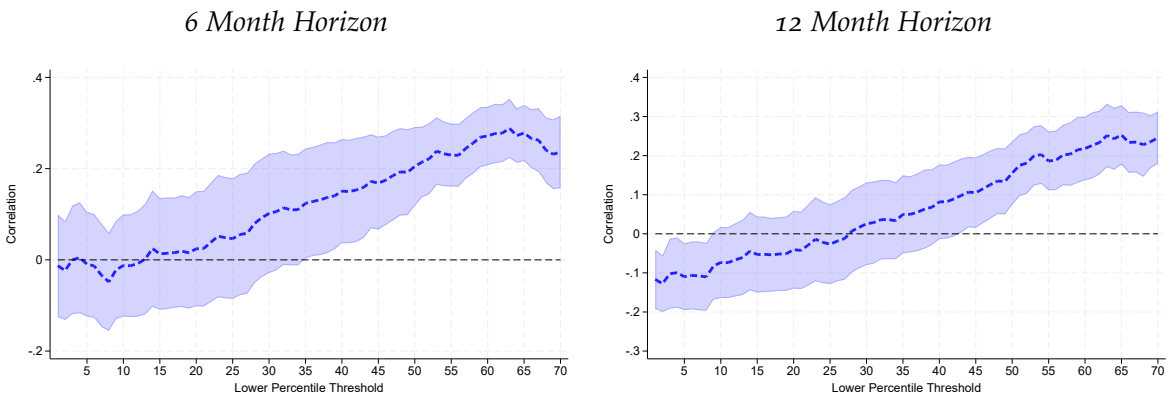
[8]The choice to retain 30 percent of the components is also motivated by the kurtosis of the import and export growth distribution, which for any given period is about 3.9 when using the Hogg (1972) kurtosis measure. For more details, please see Dolmas et al. (2005).

Figure 1: *Calibration of Trim Thresholds*

(a) *Imports*



(b) *Exports*



Notes: *These charts plot the serial correlation between current trimmed growth and aggregate import or export growth over the next 6 or 12 months based on equation (4). The horizontal axis varies the lower percentile threshold ξ_l of the trimmed index. The upper threshold ξ_h is 30 percentiles above the lower threshold. The shaded areas highlight 1 standard error confidence bands. Sample: 2002 M1 - 2024 M11.*

We illustrate our calibration in Figure 1, which plots the serial correlation between the current trimmed growth rate and the subsequent aggregate growth rate over the next six or 12 months for both imports and exports.[9] The horizontal axis indicates the lower percentile threshold ξ_l, subject to the constraint that ξ_h is 30 percentiles higher. The charts therefore display the predicative power of various cutoffs and thus different candidate trimmed growth rates. The most notable feature across all four panels is the positive relationship between the serial correlation and the size threshold up until the 65th percentile. The predictive power of the trimmed growth rate is therefore highest when it draws on subcomponents from the upper portion of the

[9]Figure A3 in the appendix provides similar charts for horizons between six and and 12 months.

import or export growth distribution, precisely because negative growth rates at the subcomponent-level are more transitory and less indicative of the aggregate growth rate than positive growth rates. This is supported by Figure A4 in the appendix, which shows that the majority of data points for aggregate import and export growth cluster around positive values.[10] Similarly, Figure 2 emphasizes that during the two major recessions in our sample, the Global Financial Crisis and the COVID-19 pandemic, import and export growth rates plummeted just to rebound shortly afterwards. In contrast, the strong growth in imports and exports following the pandemic recession lasted more than two years.[11]

In summary, we trim components below the 65th percentile ($\xi_l = 0.65$) and above the 94th percentile ($\xi_h = 0.94$) of the month-over-month import and export growth distribution. Trimming components with lower growth rates is helpful because declines in imports and exports tend to be transitory. At the same time, the trimmed growth rate also omits outliers with very high growth rates as Figure A5 in the appendix shows.

Description. Figure 2 plots monthly trimmed and aggregate nominal growth rates. The trimmed growth rates decline by less during the two recession in the sample, the Global Financial Crisis and the COVID-19 pandemic. During both episodes, imports and exports declined notably but only for a short period. On the contrary, increases in the aggregate growth rate are more persistent and the trimmed growth rate tracks elevated aggregate growth rates fairly well. Overall, the standard deviation of the trimmed growth rate is about 35 percent lower than the standard deviation of the aggregate growth rate. As discussed earlier, this reduction stems from the lower average volatility of each retained component.

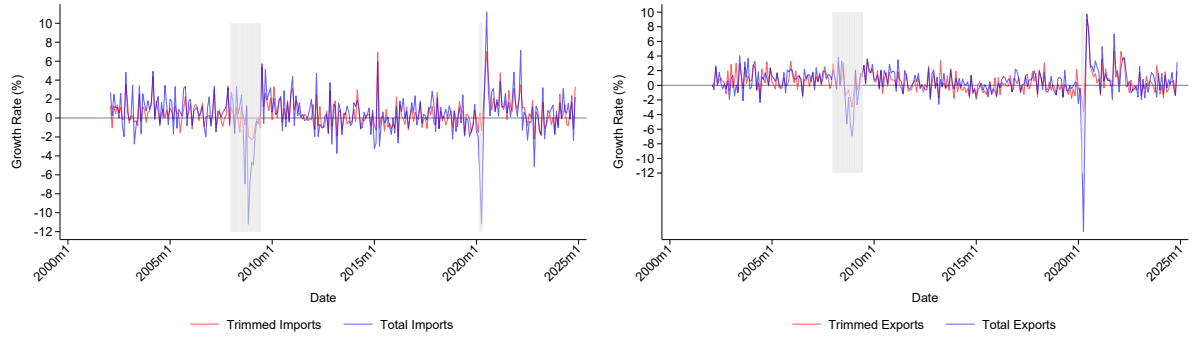
The trimmed growth rates omit 70 percent of all categories each month by construction. Figure 3 plots a histogram with the omission rate by category. The horizontal axis represents the omission rate, while the vertical axis displays the

[10] Declines in import or exports often reflect temporary port disruptions, strikes, or inventory destocking that reverse quickly. In contrast, sustained import/export increases typically reflect fundamental demand shifts, new trade relationships, or capacity expansions.

[11] Model based hyperparameter optimization using empirical Bayes also suggests $\xi_l = 0.65, \xi_h = 0.94$ to maximize the average performance of trimmed imports and exports between six and 12 month horizons.

Figure 2: *Trimmed and Aggregate Growth Rates*

(a) *Imports* ***(b)*** *Exports*

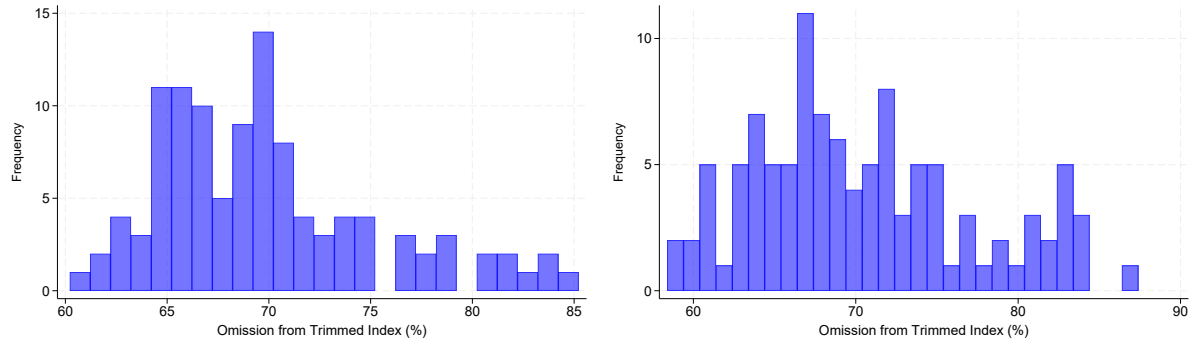


Notes: *Monthly trimmed and aggregate nominal growth rates of imports and exports in percent. The trimmed growth rate omits components below the 65th percentile and above the 94th percentile of the disaggregated import and export growth distribution. Shaded areas indicate NBER recession dates. Sample: 2002 M1 - 2024 M11.*

number of categories. The omission rate for individual categories varies between roughly 50 and 90 percent, pointing to considerable dispersion in the amount of time a component is ommitted from the trimmed series. Tables B2 and B3 in the appendix provide a more granular breakdown including a description of the subcomponents.

Figure 3: *Distribution of Omitted Subcomponents*

(a) *Imports* ***(b)*** *Exports*



Notes: *Histogram with the share of periods a category is omitted from the trimmed index for imports and exports. The trimmed growth rate omits components below the 65th percentile and above the 94th percentile of the disaggregated import and export growth distribution. Sample: 2002 M1 - 2024 M11.*

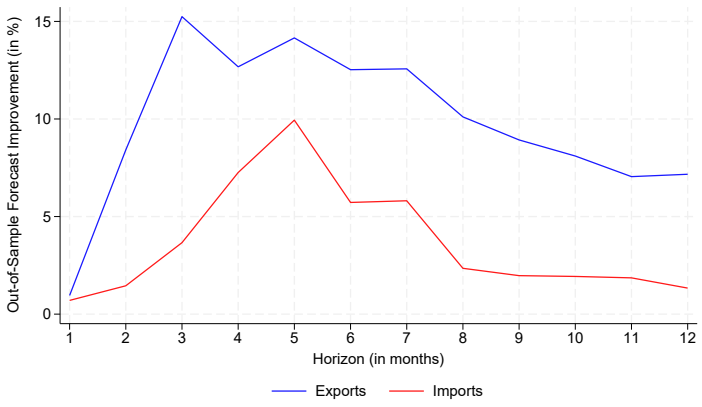
2.2. Out-Of-Sample Forecasts

This section assesses out-of-sample forecasts for aggregate imports and exports based on the trimmed growth rates, which we refer to as 'trimmed forecasts'. We show that

forecasts improve by up to 15 percent over the medium-term, relative to forecasts based on the same bivariate predictive regression but with current total import or export growth as the only independent variable.

We resort to equation (4) for out-of-sample forecasts. The independent variable is either our preferred monthly trimmed growth rate ($\xi_l = 0.65$, $\xi_h = 0.94$) or the aggregate growth rate. We initially estimate the model based on the first four years of data starting in January 2002, after which we expand the sample in one-month increments. During each iteration, we predict cumulative aggregate import or export growth $\hat{x}_{t+h|t}$ based on the latest in-sample data point for the independent variable. We then compare each out-of-sample forecast with the realization, square the residual and average the squared residuals across all iterations. This average corresponds to the out-of-sample Mean Squared Error (MSE).

Figure 4: *Out-of-Sample Forecast Improvement*



Notes*: Reduction in MSE from using the trimmed growth rates across forecasting horizons up to one year. The underlying regression is spelled out in equation (4) and estimated based on an expanding window with an initial sample of four years, starting in January 2002. The trimmed growth rate omits components below the 65th percentile and above the 94th percentile of the disaggregated import and export growth distribution. Sample: 2002 M1 - 2024 M11.*

Figure 4 highlights out-of-sample forecast improvements from trimming across forecast horizons.[12] As mentioned earlier, we choose the MSE as our criterion, which one can derive from a quadratic loss function that disproportionally penalizes outliers. The chart is based on a MSE ratio. The numerator resembles the MSE from the

[12]More precisely, we plot the percentage deterioration in out-of-sample fit from using aggregate growth instead of trimmed growth. We choose this normalization to maintain consistency across a variety of forecast comparisons in Section 3.

predictive regression with the aggregate growth rate as the independent variable. The denominator corresponds to the MSE of the trimmed forecast based on our preferred specification. We then express this ratio in percent. With this convention, the vertical axis approximates the percent forecast improvement from using the monthly trimmed growth rate instead of the aggregate growth rate as the predictor. Trimmed growth rates improve import forecasts by up to 10 percent and export forecasts by up to 15 percent. Trimmed import forecasts are most useful at horizons four to seven months ahead and trimmed export forecasts three to 12 months out. Both trimmed forecasts are not particularly useful over the near-term.

We subsequently summarize three robustness checks: The first exercise relates to the in-sample window. We prefer expanding windows over rolling windows as our default since they are less sensitive to the window length. Figure A6 shows results from a rolling four year window. This approach ignores data in the more distant past, which can be useful for detecting structural breaks. However due to the smaller sample, the estimated in-sample relationship is less precise. Because the trimmed growth rate is a less noisy predictor for future aggregate growth than aggregate growth itself (see Figure 5), this specification increases the out-of-sample fit of the trimmed forecasts relative to forecasts based on the aggregate growth rate (henceforth: 'untrimmed' forecast). Trimmed growth rates now improve import forecasts by up to 15 percent and export forecasts by more than 15 percent even at longer horizons. Trimmed exports also notably improve near-term forecasts by about 40 percent, though this is not a robust finding. Second, we choose our own seasonal adjustment for the aggregate growth rate in equation (4) to assess the predictive power of the trimmed growth rates. This provides an apples-to-apples comparison and allows us to zoom in on the effect of trimming. Figure A7 in the appendix replicates Figure 4 with the official seasonally adjusted aggregate growth rate. Because our seasonal adjustment closely follows the official seasonal adjustment, this distinction is of no practical relevance. Third, we test for significance in forecast improvements based on procedures outlined in Giacomini and White (2006) (henceforth: GW), which extend Diebold and Mariano (1995) and account for parameter uncertainty in the estimated model. The null hypothesis is that of equal conditional predictability across two models/forecasts. Results are shown

in Table 1.[13] The table shows the MSE ratio of the untrimmed forecast relative to the trimmed forecast; values greater than 1 indicate that the trimmed growth rate produced, on average, smaller errors. The stars indicate the statistical significance. Trimmed forecasts statistically outperform untrimmed forecasts across all horizons beyond three months.[14]

Table 1: *Statistical Significance: Trimmed versus Untrimmed Forecast*

Horizon	1	2	3	4	5	6	7	8	9	10	11	12
Imports	1.007**	1.014	1.037***	1.072***	1.099***	1.057***	1.058***	1.024***	1.020***	1.019***	1.018**	1.013**
Exports	1.010**	1.084	1.152	1.127***	1.141***	1.125***	1.126***	1.101***	1.089***	1.081***	1.070***	1.072***

Notes*: MSE ratios* `MSE(untrimmed)/MSE(trimmed)` *across forecasting horizons in months based on expanding windows. The regression is specified in equation (4). A ratio greater than one means that the forecast based on trimmed growth is more accurate than a forecast based on the aggregate growth rate. The trimmed growth rate omits components below the 65th percentile and above the 94th percentile of the disaggregated import and export growth distribution. Stars indicate statistical significance based on the Giacomini-White test at* $*p < 0.1$, $**p < 0.05$,$***p < 0.01$. *The null hypothesis is equal forecasting accuracy. Sample: 2002 M1 - 2024 M11.*

2.3. Intuition

In this section we show that trimmed growth rates contain a stronger and less noisy signal as to where the aggregate index is heading than the aggregate growth rate itself. The differences are larger for exports, which explains the on average superior performance of trimmed exports relative to trimmed imports as shown in Figure 4.

Figure 5 plots β^h from equation (4) over the next $h = 1, ..., 12$ months. The regression is estimated over the entire sample. The independent variable is either our preferred trimmed growth rate or the aggregate growth rate. The growth rates are standardized. β^h is therefore equivalent to the serial correlation between the independent and dependent variable. The shaded areas indicate confidence bands based on Newey-West standard errors with lags equal to the forecasting horizon.

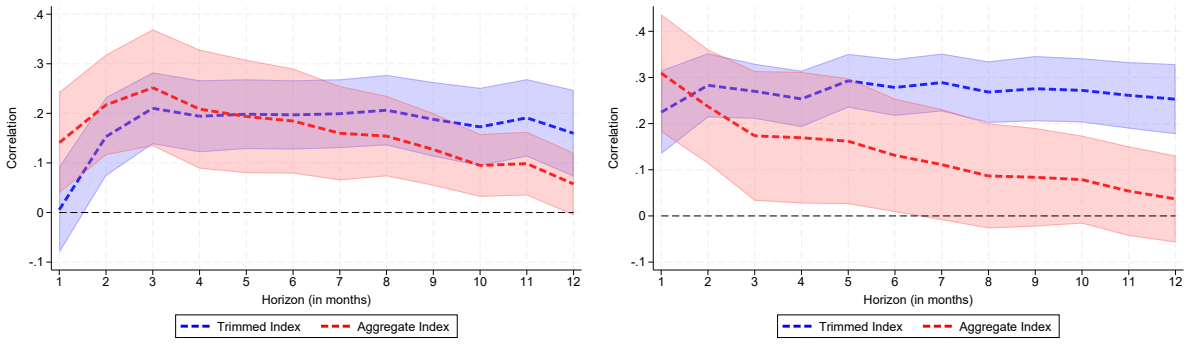
[13] The GW test examines the predictive ability of two forecasts *conditional* on the available information set. This can be implemented by regressing the loss differential, defined as the difference in squared forecast errors, on its lag and a constant. If the lagged differential has no predictive power, then one cannot distinguish both forecasts based on available information as proxied by the lagged loss differential. We estimate the regression with Newy-West (HAC) standard errors and recover a Wald statistic that is used to evaluate the null hypothesis.

[14] Table B4 in the appendix replicates Table 1 based on rolling windows. Rolling window forecasts are overall less robust, but the GW test results suggest that trimmed forecasts continue to significantly outperform forecasts based on the aggregate growth rate.

The charts have two takeaways: First, after about one quarter, trimmed growth rates are more highly correlated with future aggregate growth than the current aggregate growth rate itself. This gap widens with the forecast horizon. For example, current trimmed import growth and aggregate import growth over the subsequent 12 months is positively correlated with a correlation coefficient of about 0.15, while the correlation between current aggregate growth and future aggregate growth is only about 0.05 and not different from zero from a statistical perspective. The gap is even wider for exports. The correlation between the trimmed growth rate and aggregate growth remains between 0.25 and 0.3, while the correlation between current and future aggregate growth declines towards zero as the forecasting horizon increases. Trimmed growth rates therefore have on average more predictive power than the aggregate series itself beyond roughly one quarter. Second, the uncertainty around the correlation estimates declines when using trimmed growth as the independent variable, particularly for exports. If the correlation varies substantially over time and is therefore imprecisely estimated, one may, for example, assume a high correlation when it is in fact low. As a consequence, one would then incorrectly extract a strong signal from the current growth rate, when one should not. This in turn leads to a more inaccurate forecast.

Figure 5: *Correlation of Trimmed Growth with Future Aggregate Growth*

(a) *Imports* ***(b)*** *Exports*



Notes: *The charts plot the serial correlation between current monthly import or export growth (trimmed or aggregate) and cumulative aggregate growth over the next 1,...,12 months. The regression is specified in equation (4). The trimmed growth rate omits components below the 65th percentile and above the 94th percentile of the disaggregated import and export growth distribution. The shaded areas highlight 1 standard error confidence bands. Sample: 2002 M1 - 2024 M11.*

3. Forecast Comparison

The previous section emphasized that trimmed growth rates provide a better perspective on where aggregate imports and exports are heading than the aggregate growth rate itself. The benchmark for the comparison was a forecast based on the same model, but with the current trimmed growth rate replaced by the current aggregate growth rate. In this section, we compare trimmed forecasts to several additional forecasts. We classify these alternative forecasts into three groups: more sophisticated forecasts that rely on a richer information set or more complex statistical models; a forecast based on the weighted median as an alternative trend series; and 'naive' forecasts that exploit past trends in imports and exports. Tables 2 and 3 show that trimmed forecasts compare favorably to most of these alternatives. The notable exception is GDPNow (Atlanta Fed) which outperforms the trimmed forecast. However, GDPNow is only available for the current quarter. In contrast and as mentioned in previous sections, trimmed forecasts are most useful over the medium term.

We briefly motivate our selection for the forecast comparison. To keep this exercise tractable we focus on a series of models that cover a variety of approaches. The most sophisticated model we consider is by Koop and Korobilis (2023). They develop a Bayesian dynamic variable selection algorithm that picks the most relevant predictors across many macroeconomic, financial, and global predictors at each period in time. We adapt this time-varying parameter (TVP) model and choose a narrower set of variables to forecast imports and exports following Cook et al. (2024).[15] The model produces quarterly forecasts, so only forecasts at horizons $\{3, 6, 9, 12\}$ are shown.

The TVP model utilizes a rich information set and sophisticated statistical procedures. The next model uses a richer information set than our trimmed forecast, but a similar statistical model. In particular, we resort to monthly indicators that have been used to forecast imports and exports (see, for example, Grimme et al., 2021,

[15] Koop and Korobilis (2023) originally proposed the model to forecast inflation. We forecast a more narrowly defined set of variables, namely imports and exports. We therefore reduced the number of variables that the model is allowed to choose from to a set of variables that are more closely related to trade: GDP (BEA), disposable income (BEA), a trade weighted dollar index (FRB), trade-weighted foreign GDP (national statistical agencies and FRB), and values for imports and exports (BEA, BLS).

Lehmann, 2021) and estimate the following predictive regression for various forecast horizons h:

$$\hat{x}_{t+h|t} = \alpha^h + B^h \mathbf{Z}_t + \epsilon_{t+h}, \tag{5}$$

where $\mathbf{Z}$ is a vector that includes the following variables: capacity utilization (FRB), industrial production (FRB), new orders and new export orders (both from the ISM manufacturing index), the ISM survey on imports, the University of Michigan consumer sentiment index, the 2-year treasury yield (FRB), either import or export price indices (BLS), and either monthly nominal import or export growth rates (BEA/Census) depending on whether the dependent variable refers to imports or exports. The lag length for import and export growth is selected by the Bayesian information criterion (BIC). As in equation (4), the dependent variable $\hat{x}_{t+h|t}$ is the cumulative aggregate nominal import or export growth rate over the next h months.

The third information-rich benchmark is GDPNow (Higgins, 2014), the Atlanta Fed's continuously updated nowcast of real GDP and its subcomponents including imports and exports. Because GDPNow is updated throughout the quarter while official monthly trade data are released with a lag of two months, the average GDPNow forecast over a quarter is best compared to a three month trimmed forecast. In line with work by Faust and Wright (2009) and the widespread view that GDPNow represents a formidable near-term forecast, we expect GDPNow to outperform trimmed forecasts. The comparison is included to bound the loss of the trimmed forecast relative to the 'best-in-class' near-term forecast.

We also compare our trimmed growth rate to a weighted median. The Cleveland Fed publishes a weighted median inflation rate, and finds that this measure provides a better signal as to where the overall price index is heading. We construct the weighted median separately for imports and exports across all close to 110 subcomponents. Each subcomponent is weighted by its size. Our prior is that this alternative measure is not as useful as our trimmed measure precisely because it is best to trim components below the 65th percentile.

The remaining three benchmarks are based on 'rules of thumb'. We consider a

three-month moving-average (MA3) and replace the trimmed growth rate in equation (4) with a moving average of the aggregate monthly growth rate:

$$\hat{x}_{t+h|t} = \alpha^h + \beta^h \left(\frac{\hat{x}_{t|t-1} + \hat{x}_{t-1|t-2} + \hat{x}_{t-2|t-3}}{3} \right) + \epsilon_{t+h}. \tag{6}$$

A comparison of trimmed forecasts with moving-average forecasts addresses whether the trimmed growth rate is just a smoother version of the aggregate growth rate due to the omission of outliers. We also consider a historic average (Hist. Avg.) for which the forecast for say aggregate import growth over the next 4 months corresponds to the average 4 month import growth in the sample up to the observation that is forecasted. In other words, this forecast is based on the above regression but without an independent variable. Last but not least, we consider a no-change forecast, which is based on the premise that the best forecast for the next h months is the growth rate over the past h months. This would be the case if the growth rate followed a random walk. Overall, these three benchmarks represent different levels of sensitivity to recent data: the no-change forecast is most sensitive, since it propagates the most recent observation directly; the historic average is least sensitive; and the MA3 forecast falls in between, smoothing recent observations.

Table 2: *Forecast Comparisons: Imports*

Horizon	1	2	3	4	5	6	7	8	9	10	11	12
TVP			1.141***			1.267***			1.282***			1.142***
Indicator	0.825	1.008	1.364	1.473	1.683*	1.736*	1.791**	1.728***	1.852***	1.931***	1.952***	2.642**
GDPNow			0.490***									
Median	1.001	1.021***	1.043***	1.036***	1.041**	1.030**	1.029***	1.017***	1.013***	1.010	1.013	0.999
MA3	0.963	1.099***	1.208***	1.208***	1.234***	1.116***	1.088***	1.070***	1.033***	1.036***	1.038***	1.034***
Hist. Avg.	0.981	1.004	1.030	1.017	1.026	1.020	1.026	1.035	1.017	1.013	1.026	1.004
NoChange	1.585**	1.398	1.460***	1.554***	1.637***	1.736***	1.877***	1.995***	2.025***	2.068***	2.132***	2.127***

Notes*: MSE ratios* `MSE(model)/MSE(trimmed)` *across forecasting horizons in months based on expanding windows. See the text for details on the models. A ratio greater than one means that the forecast based on trimmed growth is more accurate. The trimmed growth rate omits components below the 65th percentile and above the 94th percentile of the disaggregated import and export growth distribution. Stars indicate statistical significance based on the Giacomini-White test at * $p < 0.1$, **$p < 0.05$,***$p < 0.01$. The null hypothesis is equal forecasting accuracy. Sample: 2002 M1 - 2024 M11.*

Tables 2 and 3 show the MSE of these alternative models relative to the trimmed forecast. A value greater than one indicates an on-average smaller MSE for the trimmed forecast. As in previous exercises, we evaluate these forecasts out-of-sample based

on expanding windows, with an initial four year window starting in January 2002.[16] The star symbols indicate statistical significance based on the GW test introduced in Section 2.2. In each case we contrast the alternative forecast with the trimmed forecast. The null hypothesis of this test reflects equal predictability across both forecasts.

The tables have the following takeaways: Trimmed forecasts are competitive with information-rich benchmarks beyond the current quarter. Against the indicator model, the trimmed forecast outperforms at every horizon beyond four months. The MSE ratio peaks at over two, suggesting that the MSE of the indicator model can be twice as large.[17] The trimmed forecasts also outperform the TVP model by up to 30 percent. However, GDPNow – our benchmark for the best near-term forecast – is about 50 percent more accurate than the trimmed forecast. Last but not least, the trimmed forecasts outperform forecasts based on the weighted median and naive rule-of-thumb forecasts with the exception of the historic average and imports. Overall, we conclude that trimmed growth rates are a useful predictor over the medium term.[18]

Table 3: *Forecast Comparisons: Exports*

Horizon	1	2	3	4	5	6	7	8	9	10	11	12
TVP			1.144***			1.306***			1.337***			1.101
Indicator	1.197	1.071	1.318	1.373*	1.436**	1.511**	1.610**	1.603***	1.790***	1.914***	1.752***	1.969***
GDPNow			0.442*									
Median	0.997	1.083	1.114	1.121	1.135**	1.106***	1.107***	1.082***	1.082***	1.070***	1.074***	1.067***
MA3	1.132**	1.353*	1.348***	1.273***	1.260***	1.233***	1.204***	1.153***	1.138***	1.126***	1.113***	1.098***
Hist. Avg.	1.041	1.070	1.061	1.047	1.076*	1.079*	1.086**	1.077***	1.080***	1.077***	1.068***	1.060***
NoChange	1.386	1.772	1.838***	1.828***	1.947***	2.033***	2.124***	2.175***	2.165***	2.184***	2.185***	2.172***

Notes*: MSE ratios* `MSE(model)/MSE(trimmed)` *across forecasting horizons in months based on expanding windows. See the text for details on the models. A ratio greater than one means that the forecast based on trimmed growth is more accurate. The trimmed growth rate omits components below the 65th percentile and above the 94th percentile of the disaggregated import and export growth distribution. Stars indicate statistical significance based on the Giacomini-White test at* * $p < 0.1$, **$p < 0.05$,***$p < 0.01$. *The null hypothesis is equal forecasting accuracy. Sample: 2002 M1 - 2024 M11.*

[16]The TVP model is estimated over the entire sample, as it is time intensive to estimate and by construction supposed to filter out the most relevant predictors at each point in time.

[17]Though not displayed, we also examined each indicator separately. The results are unchanged. The only exception is capacity utilization, which outperforms trimmed growth during the COVID-19 supply shock.

[18]Tables B5 and B6 in the appendix confirm these findings based on expanding windows.

4. When Are Trimmed Forecasts Useful?

We previously showed that simple bivariate predictive regressions based on the trimmed growth rate improve forecasts for aggregate imports and exports *on average*. In this section, we explore *when* trimmed forecasts are useful. While trimmed forecasts are useful throughout the sample period, forecast improvements are largest during periods of heightened volatility with large but temporary declines in imports and exports.

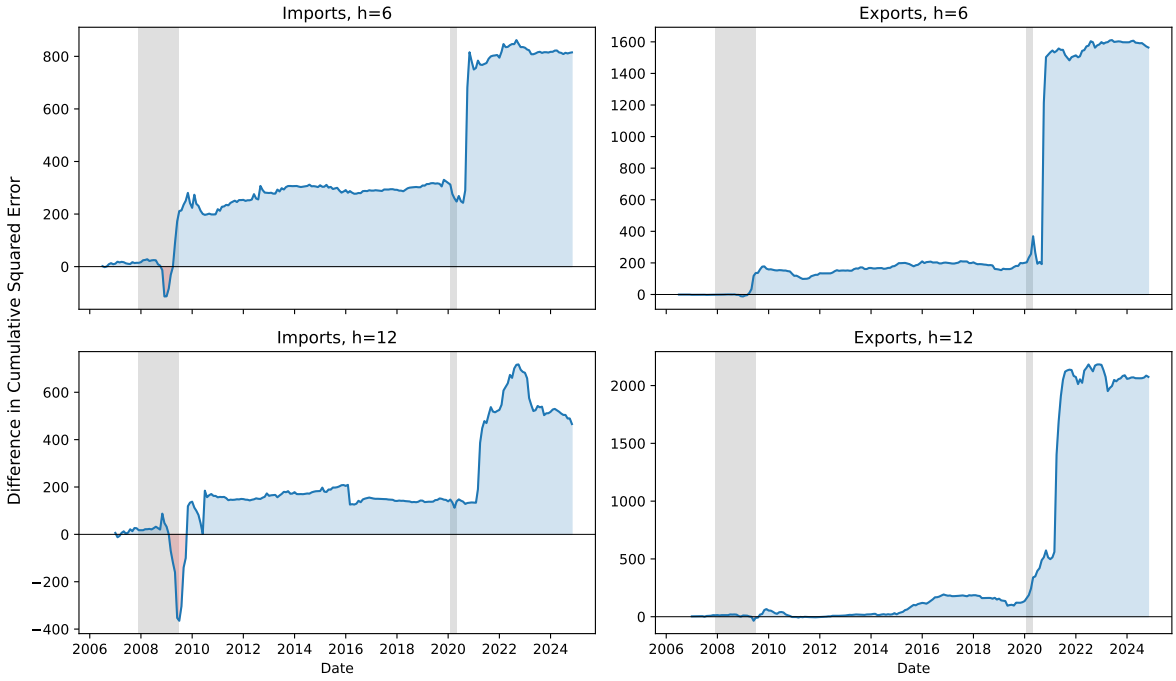
Figure 6 shows the difference in the cumulative squared error between the untrimmed and trimmed forecast over time. For illustrative purposes we focus on 6 and twelve month horizons. To be precise, the figure shows the following object, which we define as Δ_t:

$$\Delta_t \equiv \sum_{\tau \leq t} \left((\hat{\epsilon}_{\tau+h})^2 - \left(\hat{\epsilon}^{TM}_{\tau+h} \right)^2 \right), \tag{7}$$

where $\hat{\epsilon}^{TM}_{\tau+h}$ is the realized error from regression (4) and $\hat{\epsilon}_{\tau+h}$ the corresponding error from the same regression with the trimmed growth rate replaced by the aggregate growth rate. The forecasts are based on expanding windows with an initial window size of four years. The level of Δ_t is indicative of the overall relative performance of the two forecasts up until year-month t. A positive value is equivalent to a superior performance of the trimmed forecast. The slope of Δ_t emphasizes the relative performance of the forecasts at a particular time. If the slope is positive, the trimmed forecast contemporaneously outperforms the untrimmed forecast.

The trimmed forecast is more precise than the untrimmed forecast during COVID-19 as shown in Figure 6. During COVID-19, imports and exports declined only temporarily. The negative aggregate growth rate was therefore not a good predictor as to where aggregate imports or exports were heading (see Figure 2). The trimmed growth rate in contrast discounts the temporary drop, which improves the trimmed forecast primarily during the pandemic recovery. Trimmed growth rates were also useful during the Global Financial Crisis (GFC). However, the decline imports and exports during the GFC was more persistent; accordingly the trimmed and aggregate

Figure 6: *Forecast Accuracy over Time: Squared Errors*

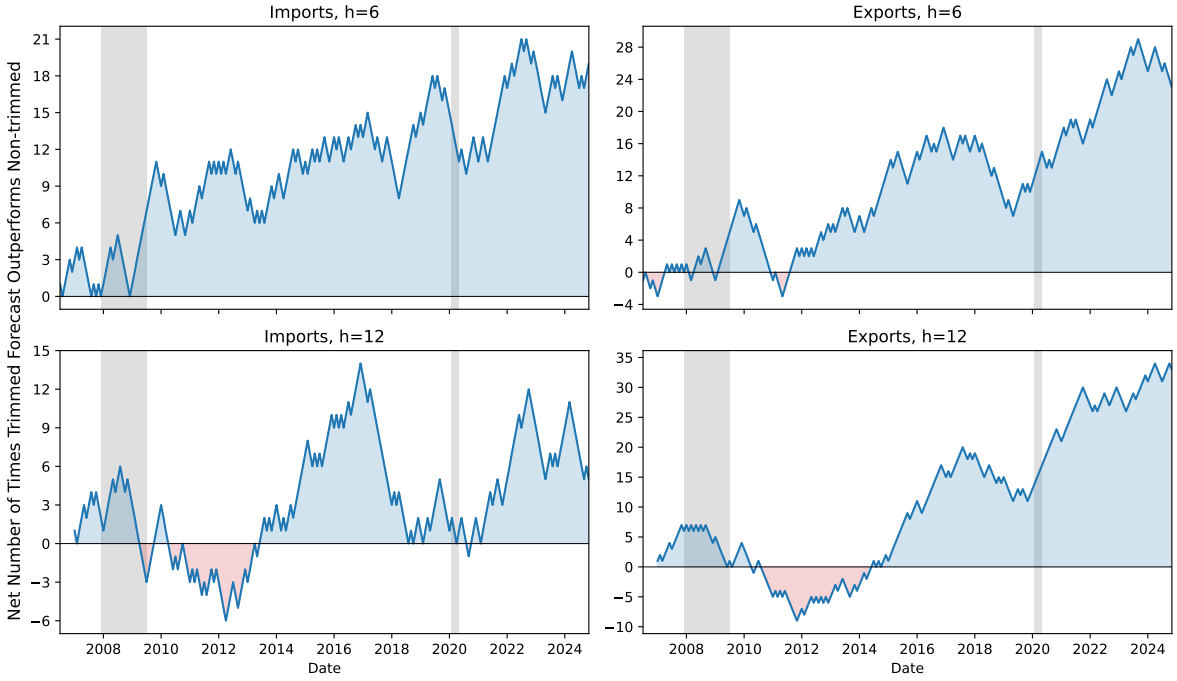


Notes*: Cumulative difference in squared errors across time between the trimmed and untrimmed forecast. The underlying regression is spelled out in equation (4) and estimated based on an expanding window with an initial sample of four years, starting in January 2002. The trimmed growth rate omits components below the 65th percentile and above the 94th percentile of the disaggregated import and export growth distribution. Sample: 2002 M1 - 2024 M11.*

growth rates differ by less and improvements in forecasting accuracy are more limited. Last but not least, we want to highlight that trimmed growth rates are also useful during 'normal' times. This can be inferred from Figure 7, which plots the cumulative sum of an indicator that is +1, if the error of the trimmed forecast is smaller than the error of the untrimmed forecast in absolute values. Otherwise, the indicator is -1. For the six month horizon, the trimmed forecast outperforms the untrimmed forecast for most of the sample period. In contrast, at the 12 month horizon, the trimmed forecast underperformed between 2010 to 2012 and then in 2018. Quantitatively, however, trimmed forecasts are more relevant when imports and exports are more volatile.[19]

[19]Figures A8 to A19 in the appendix extend this comparison to the alternative forecasts from Section 3. The overall findings are similar.

Figure 7: *Forecast Accuracy over Time: Better or Worse*



Notes*: Cumulative sum of an indicator that is +1, if the error of the trimmed forecast is smaller than the error of the untrimmed forecast in absolute terms. Otherwise, the indicator is -1. The underlying regression is spelled out in equation (4) and estimated based on an expanding window with an initial sample of four years, starting in January 2002. The trimmed growth rate omits components below the 65th percentile and above the 94th percentile of the disaggregated import and export growth distribution. Sample: 2002 M1 - 2024 M11.*

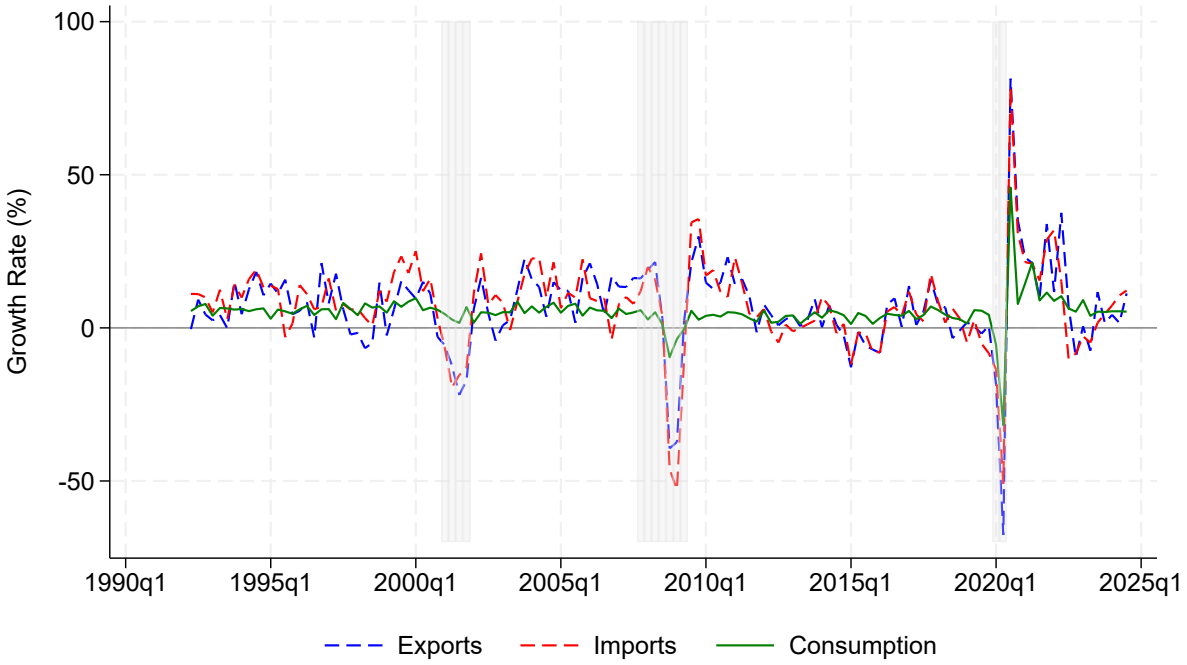
5. Conclusion

This paper introduces monthly trimmed import and export growth rates that provide a better perspective on underlying trade dynamics than the aggregate growth rates themselves. We then show that simple bivariate predictions based on the trimmed growth rates outperform several other forecasts out-of-sample particularly over four to 12 month horizons. We want to emphasize that these forecast comparisons are only illustrative. In fact, one might be able to improve trade forecasts by combining trimmed growth rates with other predictors.

We regularly update the trimmed growth rates. We hope they are useful for practitioners who evaluate trade dynamics and beyond. For example, a more reliable view on the trade balance may inform policymakers about the underlying strength of the domestic economy.

A. Appendix: Figures

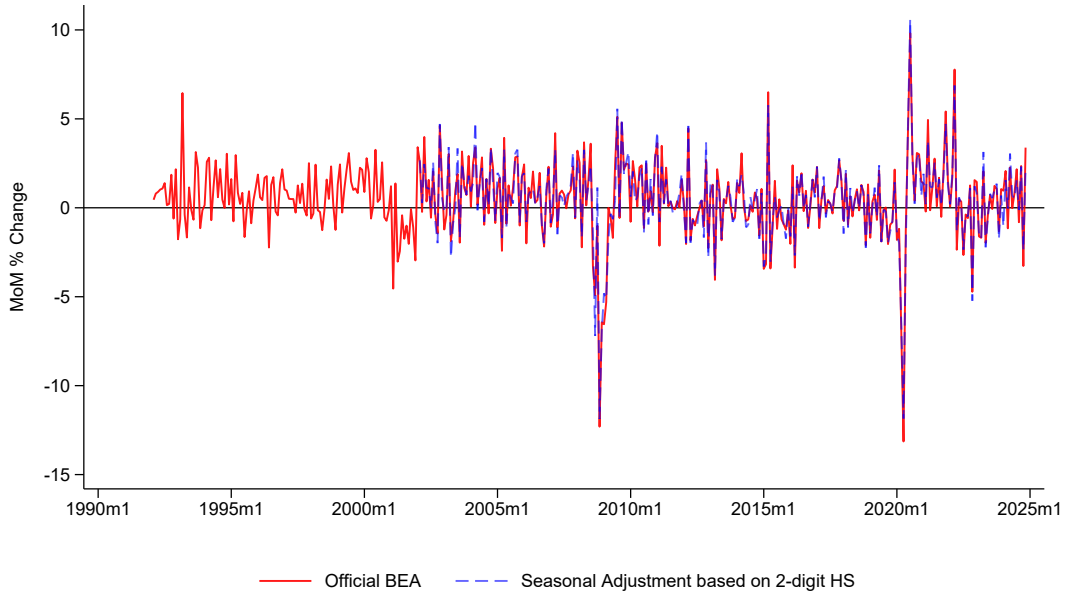
Figure A1: *Growth Rates: Imports, Exports and Consumption*



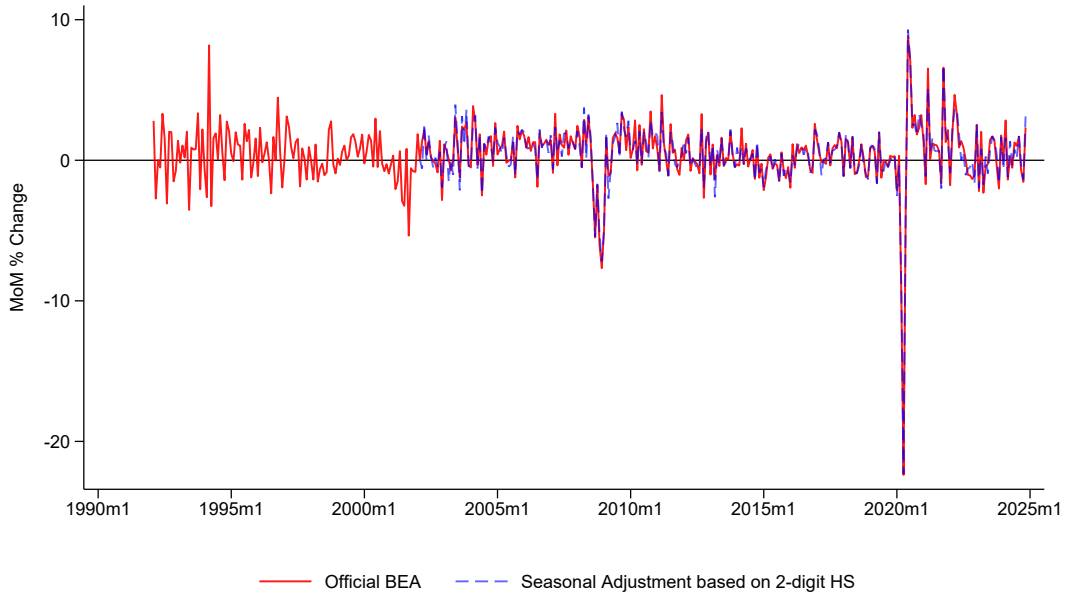
Notes*: Annualized quarterly nominal growth rates of imports, exports and consumption in percent. Shaded areas indicate NBER recession dates. Sample: 1992 Q2 - 2024 Q3.*

Figure A2: *Seasonal Adjustments*

(a) *Imports*

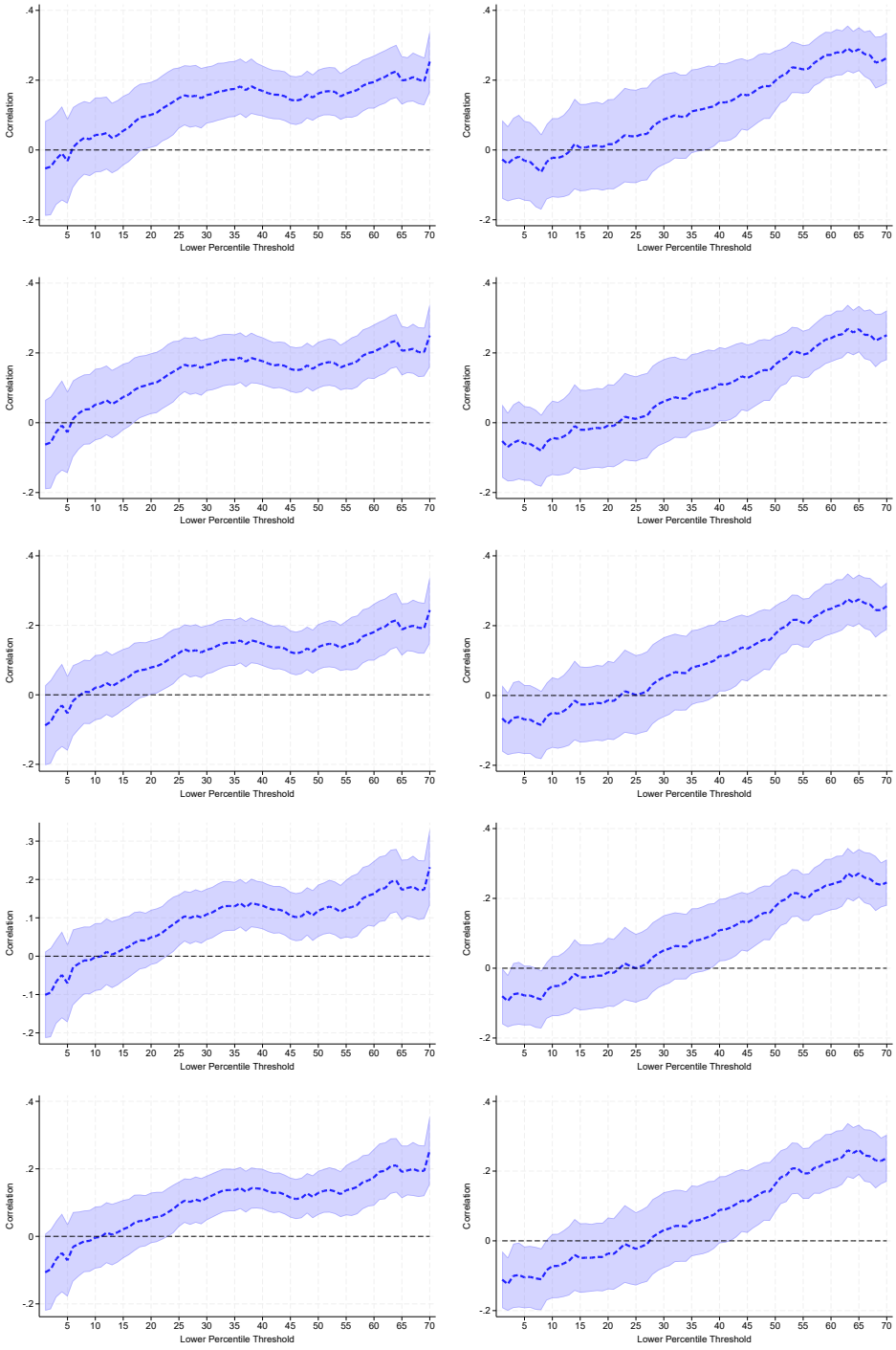


(b) *Exports*



Notes: *Seasonally adjusted monthly aggregate nominal import and export growth rates based on the official BEA release and our own seasonal adjustment at the 2-digit HS level.*

Figure A3: *Calibration of Trim Thresholds*



Notes: *These charts plot the serial correlation between current trimmed growth and aggregate import or export growth over the next 7 to 11 months based on equation (4). Each row corresponds to a different horizon starting at 7 months. The first column refers to imports the second column to exports. The horizontal axis varies the lower percentile threshold ξ_l of the trimmed index. The upper threshold ξ_h is 30 percentiles above the lower threshold. The shaded areas highlight 1 standard error confidence bands. Sample: 2002 M1 - 2024 M11.*

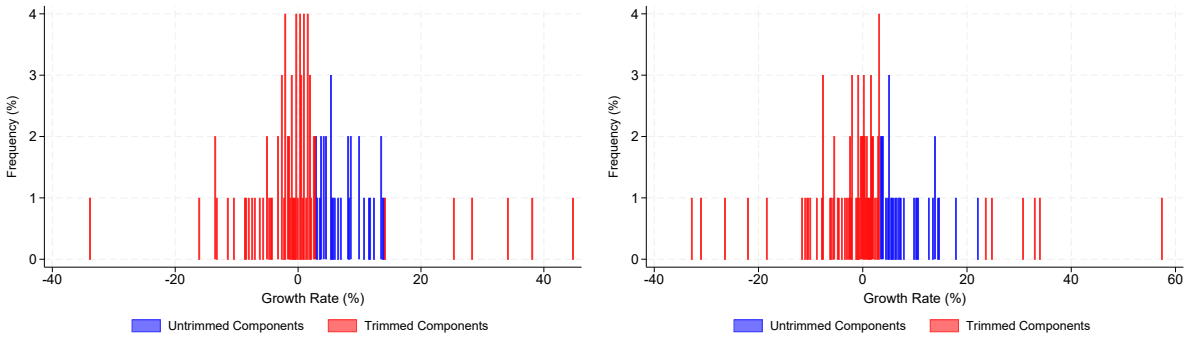
Figure A4: *Aggregate Import and Export Growth*



Notes*: Histogram with the annualized quarterly nominal growth rates of imports and exports in percent. Solid lines represent kernel densities and bars the raw data. Sample: 1992 Q2 - 2024 Q3.*

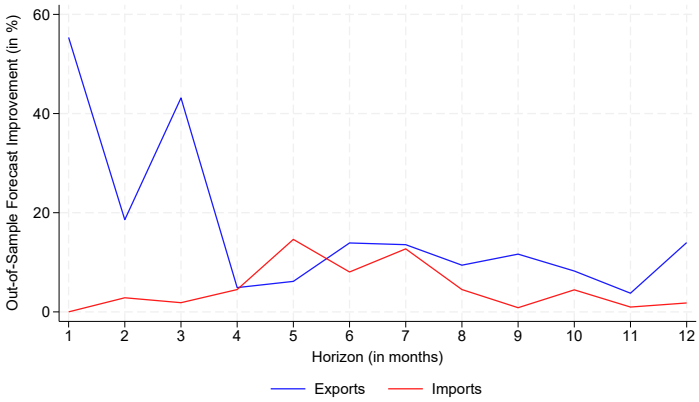
Figure A5: *Growth Rates of Subcomponents*

(a) *Imports* ***(b)*** *Exports*



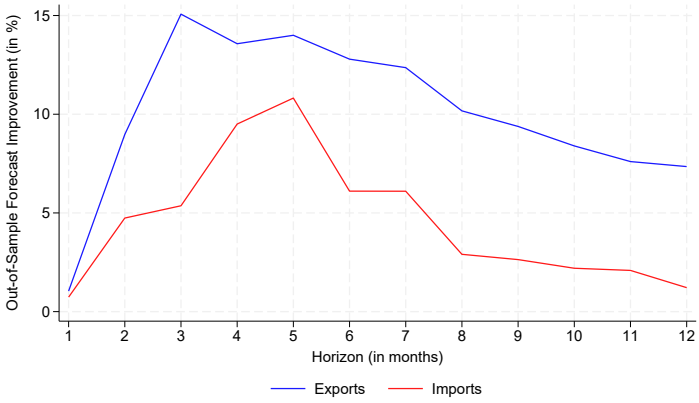
Notes: *Histogram with the monthly nominal growth rate across subcomponents for imports and exports. The chart distinguishes between untrimmed (blue) and trimmed (red) components. The chart is based on a snapshot from 2021 M1, though similar patterns hold for other months as well.*

Figure A6: *Out-of-Sample Forecast Improvement: Rolling Windows*



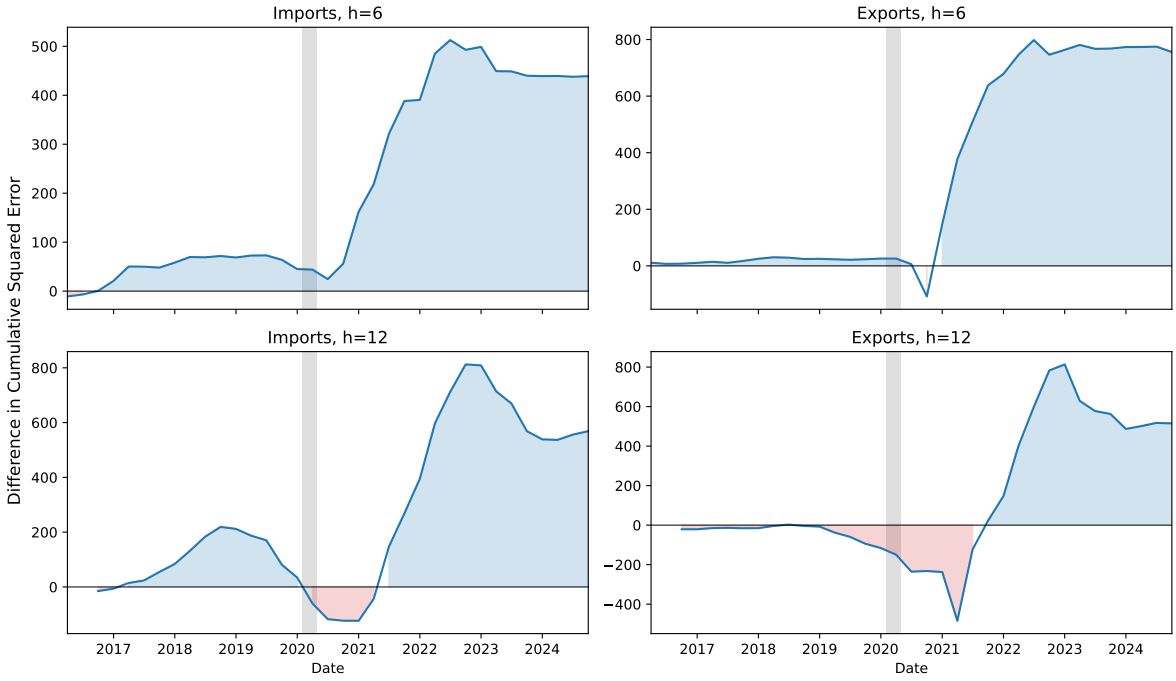
Notes*: Reduction in MSE from using the trimmed growth rates across forecasting horizons up to one year. The underlying regression is spelled out in equation (4) and estimated based on rolling windows of fours years, starting in January 2002. The trimmed growth rate omits components below the 65th percentile and above the 94th percentile of the disaggregated import and export growth distribution. Sample: 2002 M1 - 2024 M11.*

Figure A7: *Out-of-Sample Forecast Improvement: Official Data*



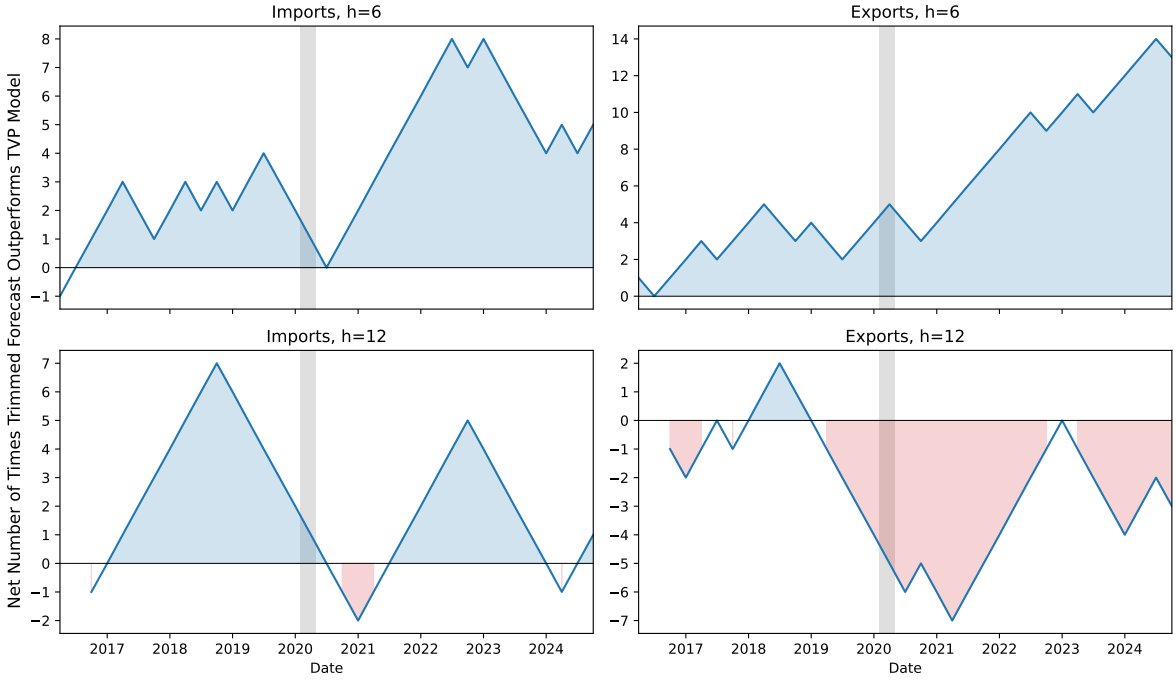
Notes*: Reduction in MSE from using the trimmed growth rates across forecasting horizons up to one year. The aggregate growth rate is based on the official BEA release and not our own seasonal adjustment. The underlying regression is spelled out in equation (4) and estimated based on an expanding window with an initial sample of four years, starting in January 2002. The trimmed growth rate omits components below the 65th percentile and above the 94th percentile of the disaggregated import and export growth distribution. Sample: 2002 M1 - 2024 M11.*

Figure A8: *Forecast Accuracy over Time Relative to TVP Benchmark: Squared Errors*



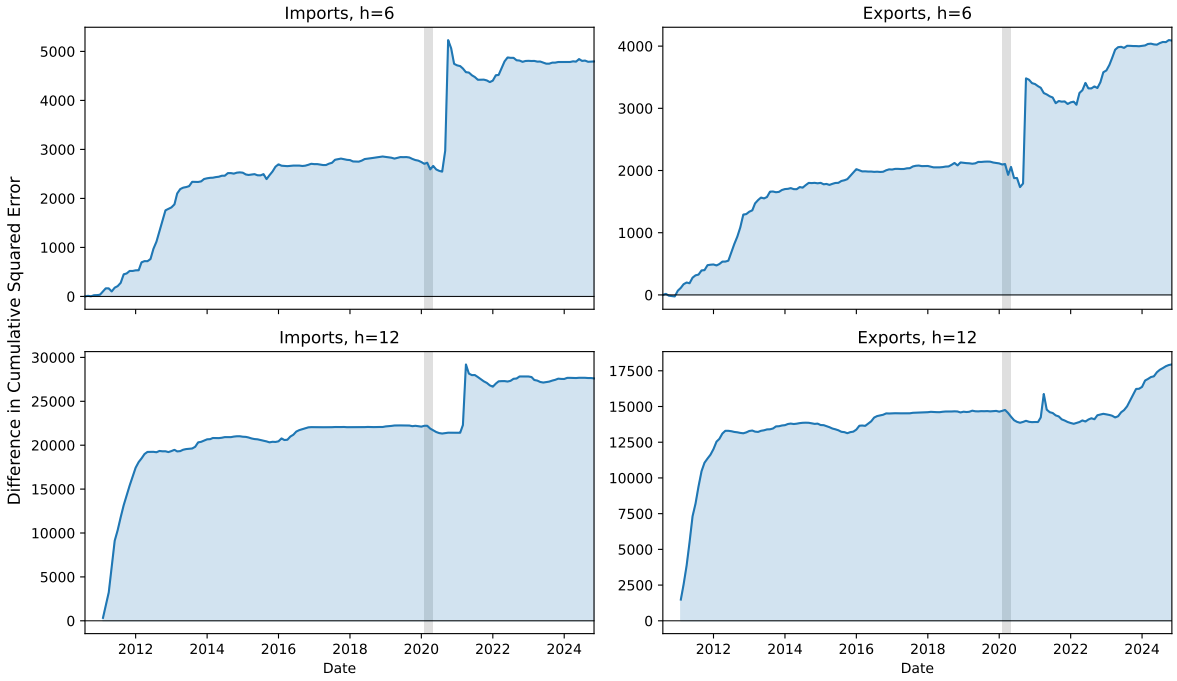
Notes*: Cumulative difference in squared errors across time between the trimmed and TVP forecast. Sample: 2002 M1 - 2024 M11.*

Figure A9: *Forecast Accuracy over Time Relative to TVP Benchmark: Better or Worse*



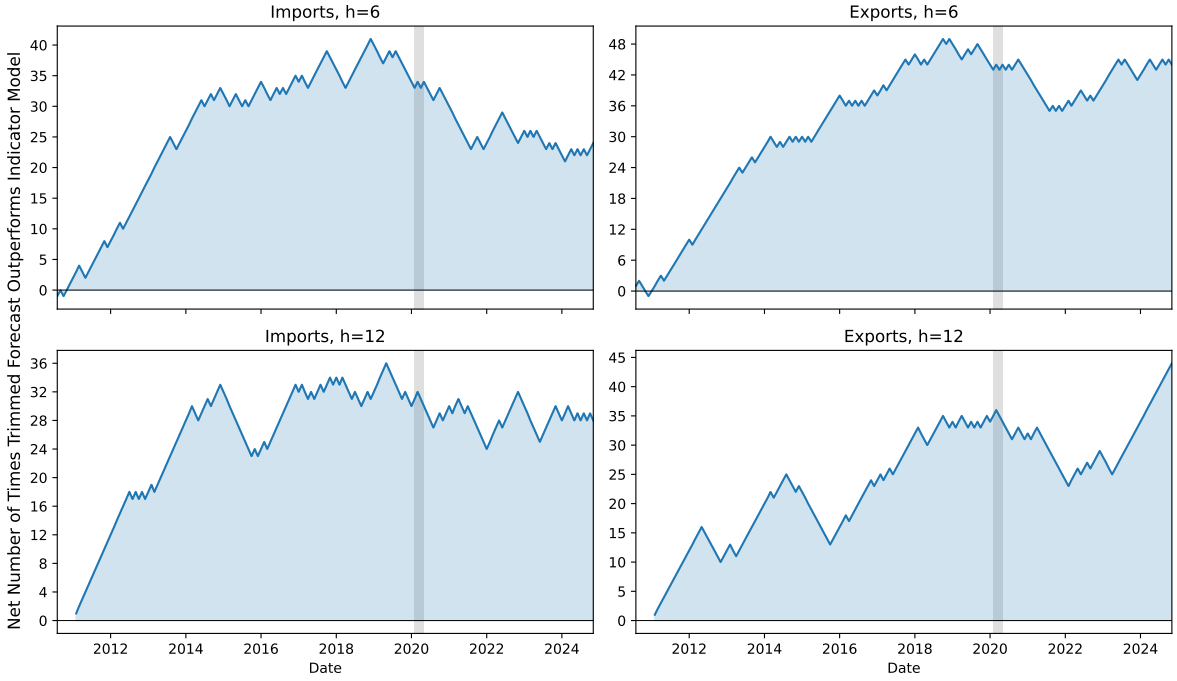
Notes*: Cumulative sum of an indicator that is +1, if the error of the trimmed forecast is smaller than the error of the TVP forecast in absolute terms. Otherwise, the indicator is -1. Sample: 2002 M1 - 2024 M11.*

Figure A10: *Forecast Accuracy over Time Relative to Indicator Benchmark: Squared Errors*



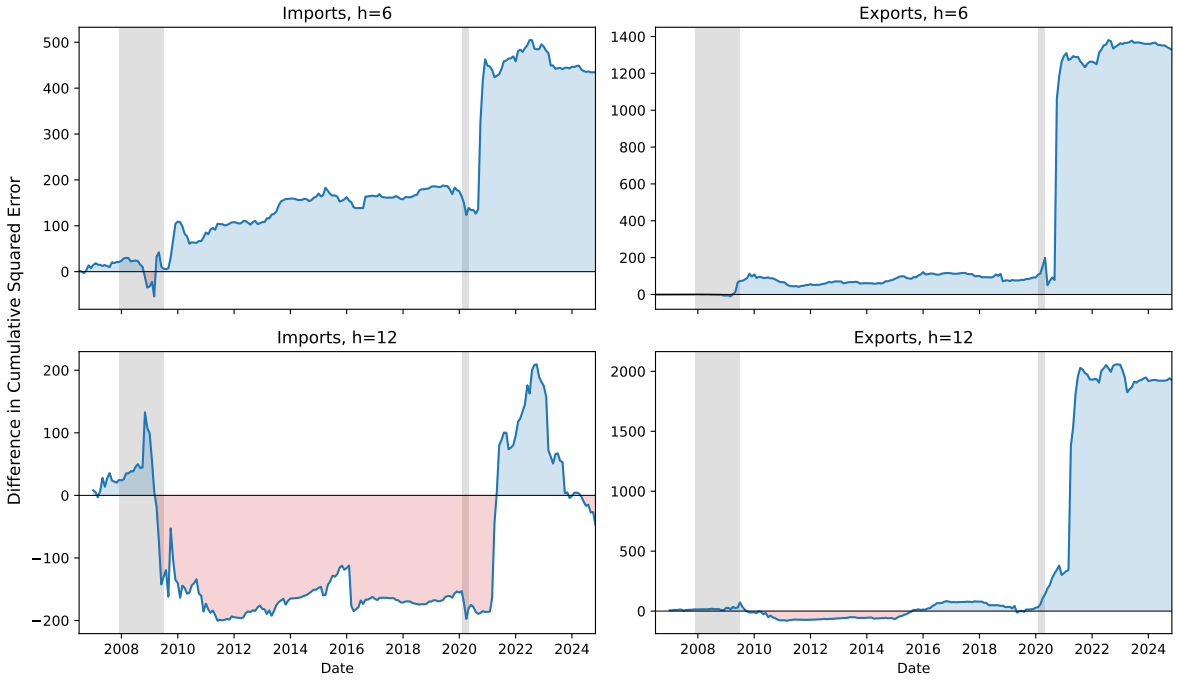
Notes*: Cumulative difference in squared errors across time between the trimmed and indicator forecast. Sample: 2002 M1 - 2024 M11.*

Figure A11: *Forecast Accuracy over Time Relative to Indicator Benchmark: Better or Worse*



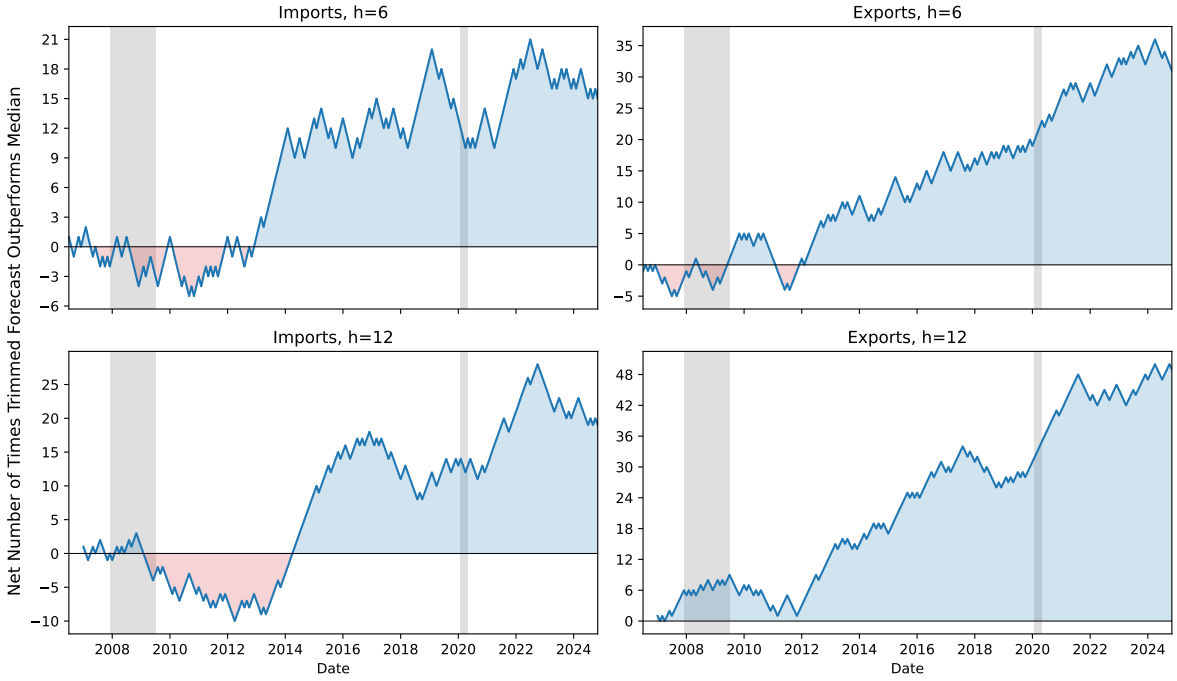
Notes*: Cumulative sum of an indicator that is +1, if the error of the trimmed forecast is smaller than the error of the indicator forecast in absolute terms. Otherwise, the indicator is -1. Sample: 2002 M1 - 2024 M11.*

Figure A12: *Forecast Accuracy over Time Relative to Weighted Median Benchmark: Squared Errors*



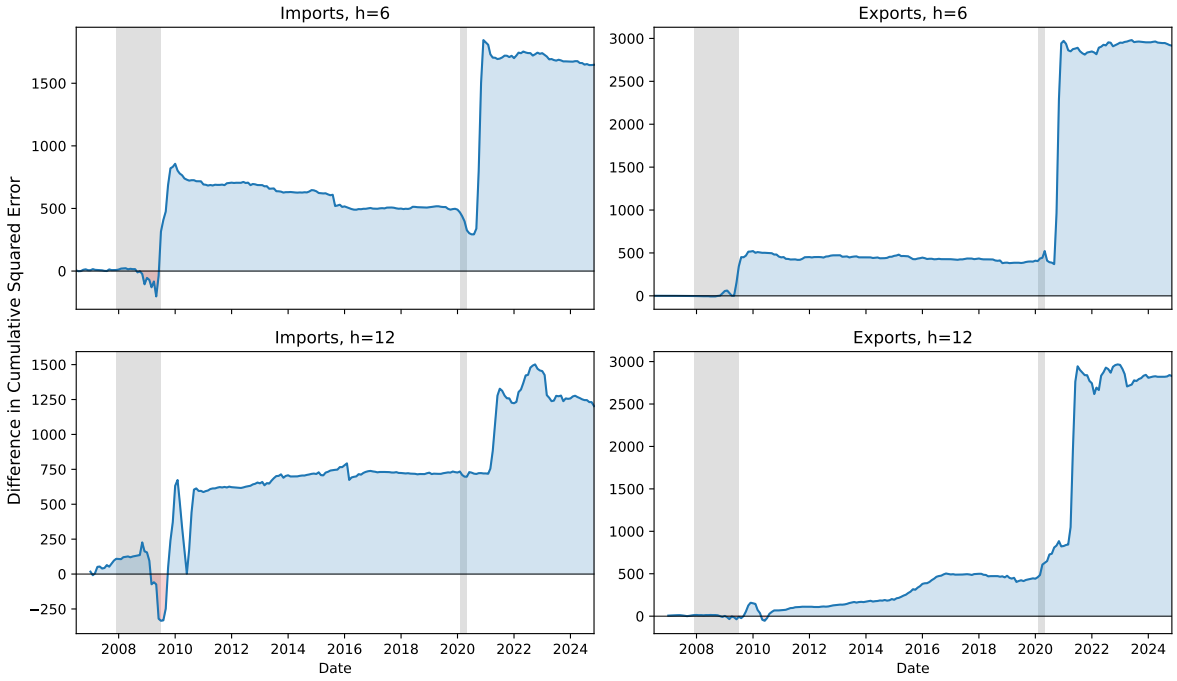
Notes*: Cumulative difference in squared errors across time between the trimmed and weighted median forecast. Sample: 2002 M1 - 2024 M11.*

Figure A13: *Forecast Accuracy over Time Relative to Weighted Median Benchmark: Better or Worse*



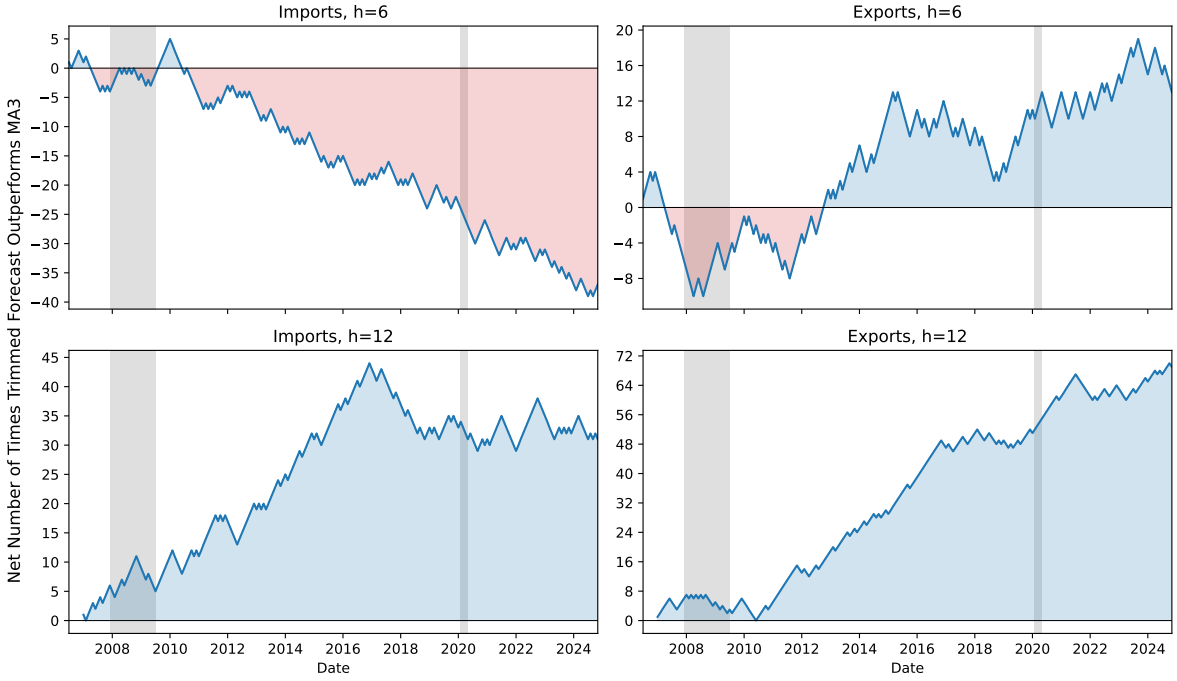
Notes*: Cumulative sum of an indicator that is +1, if the error of the trimmed forecast is smaller than the error of the weighted median forecast in absolute terms. Otherwise, the indicator is -1. Sample: 2002 M1 - 2024 M11.*

Figure A14: *Forecast Accuracy over Time Relative to MA3 Benchmark: Squared Errors*



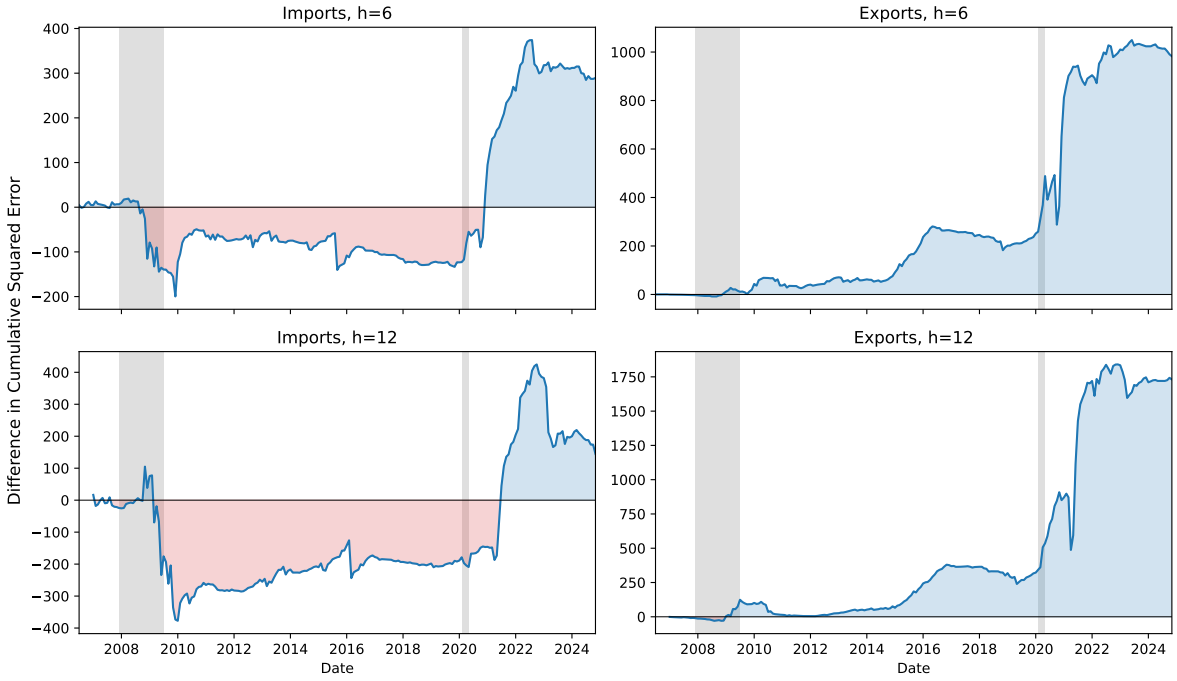
Notes*: Cumulative difference in squared errors across time between the trimmed and MA3 forecast. Sample: 2002 M1 - 2024 M11.*

Figure A15: *Forecast Accuracy over Time Relative to MA3 Benchmark: Better or Worse*



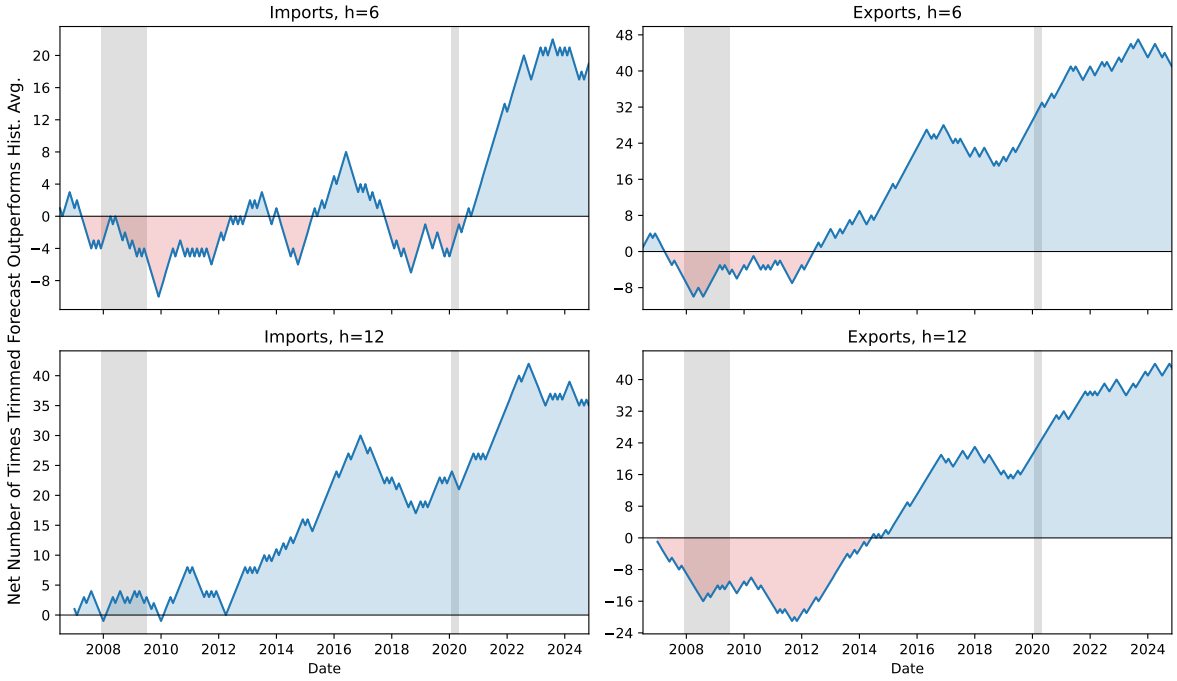
Notes*: Cumulative sum of an indicator that is +1, if the error of the trimmed forecast is smaller than the error of the MA3 forecast in absolute terms. Otherwise, the indicator is -1. Sample: 2002 M1 - 2024 M11.*

Figure A16: *Forecast Accuracy over Time Relative to Historical Average Benchmark: Squared Errors*



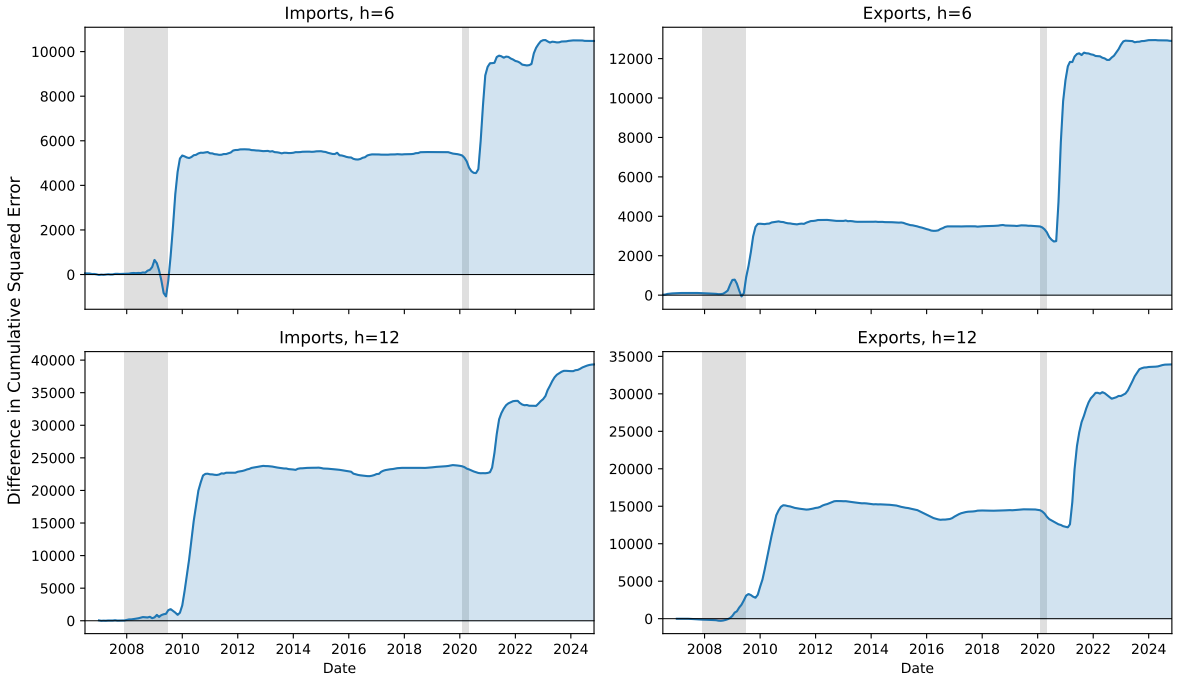
Notes*: Cumulative difference in squared errors across time between the trimmed and historical average forecast. Sample: 2002 M1 - 2024 M11.*

Figure A17: *Forecast Accuracy over Time Relative to Historical Average Benchmark: Better or Worse*



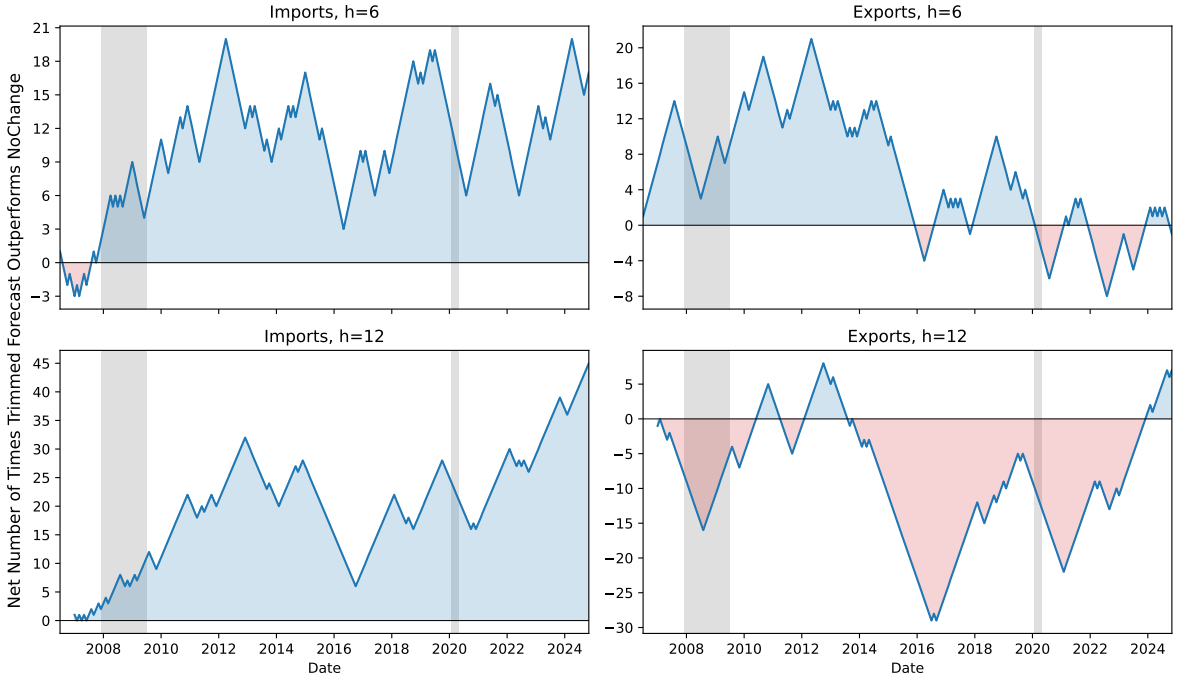
Notes*: Cumulative sum of an indicator that is +1, if the error of the trimmed forecast is smaller than the error of the historical average in absolute terms. Otherwise, the indicator is -1. Sample: 2002 M1 - 2024 M11.*

Figure A18: *Forecast Accuracy over Time Relative to No-Change Benchmark: Squared Errors*



Notes*: Cumulative difference in squared errors across time between the trimmed and no-change forecast. Sample: 2002 M1 - 2024 M11.*

Figure A19: *Forecast Accuracy over Time Relative to No-Change Benchmark: Better or Worse*



Notes*: Cumulative sum of an indicator that is +1, if the error of the trimmed forecast is smaller than the error of the no-change forecast in absolute terms. Otherwise, the indicator is -1. Sample: 2002 M1 - 2024 M11.*

B. Appendix: Tables

Table B1: *Trade Forecast Errors*

Horizon (Quarters)	Imports	Exports
$h = 0$	5.56	5.03
$h = 1$	44.54	19.86

Notes*: The table shows the MSE of import and export forecasts relative to the MSE of the current quarter consumption forecast. All forecasts are from GDPNow and expressed in annualized quarter-over-quarter real growth rates in percent. GDPNow automatically switches to the next quarter at the end of the current quarter. We treat the latest forecast within a quarter as the current-quarter forecast ($h = 0$). For the one-quarter-ahead forecast ($h = 1$), we wait for the first update after the last monthly import/export release from the previous quarter. Sample: 2014 Q2 - 2024 Q4. We exclude the Covid-19 period from 2020 Q1 - 2020 Q4 during which forecast errors for imports and exports were much larger.*

Table B2: *Omitted Subcomponents: Imports*

Category	HS Code (Goods Only)	Omission Rate (%)
Machinery	84	84.31
Ships	89	83.94
Government Goods and Services	-	83.58
Paper	48	83.21
Plastics	39	82.12
Art	97	81.75
Lead	78	81.02
Tin	80	80.66
Glass	70	78.83
Other Business Services	-	78.83
Cocoa	18	63.87
Dairy & Eggs	4	63.87
Cosmetics & Perfumes	33	63.5
Food Preparations	21	63.14
Albuminoids & Enzymes	35	63.14
Meat & Fish Preparations	16	62.77
Salt, Sulfur, Stone	25	62.41
Mineral Fuels	27	62.04
Food Residues & Animal Feed	23	61.31
Coffee & Spices	9	60.22

Notes*: The table shows subcomponents with the highest and lowest omission rates. The trimmed growth rate omits components below the 65th percentile and above the 94th percentile of the disaggregated import and export growth distribution. Sample: 2002 M1 - 2024 M11.*

Table B3: *Omitted Subcomponents: Exports*

Category	HS Code (Goods Only)	Omission Rate (%)
Furskins	43	87.23
Silk	50	84.31
Paper	48	84.31
Precision Instruments	90	83.58
Electrical Equipment	85	83.21
Vegetable Materials	14	82.85
Arms & Ammunition	93	82.85
Plastics	39	82.85
Use of Intellectual Property	-	82.48
Machinery	84	82.12
Clocks & Watches	91	61.68
Musical Instruments	92	61.31
Construction Services	-	61.31
Coffee & Spices	9	61.31
Cocoa	18	61.31
Ceramics	69	60.95
Nickel	75	60.22
Copper	74	60.22
Mineral Fuels	27	58.76
Maintenance and Repair Services	-	58.39

Notes*: The table shows subcomponents with the highest and lowest omission rates. The trimmed growth rate omits components below the 65th percentile and above the 94th percentile of the disaggregated import and export growth distribution. Sample: 2002 M1 - 2024 M11.*

Table B4: *Rolling Window Statistical Significance: Trimmed versus Trimmed Forecast*

Horizon	1	2	3	4	5	6	7	8	9	10	11	12
Imports	0.988	1.028***	1.018***	1.045	1.146***	1.081***	1.127***	1.045**	1.008***	1.044***	1.010	1.018**
Exports	1.554***	1.186***	1.432***	1.049	1.062***	1.139***	1.136***	1.094**	1.117***	1.082***	1.038	1.140***

Notes*: MSE ratios* `MSE(untrimmed)/MSE(trimmed)` *across forecasting horizons in months based on four year rolling windows. The regression is specified in equation (4). A ratio greater than one means that the forecast based on trimmed growth is more accurate than a forecast based on the aggregate growth rate. The trimmed growth rate omits components below the 65th percentile and above the 94th percentile of the disaggregated import and export growth distribution. Stars indicate statistical significance based on the Giacomini-White test at* $*p < 0.1$, $**p < 0.05$,$***p < 0.01$. *The null hypothesis is equal forecasting accuracy. Sample: 2002 M1 - 2024 M11.*

Table B5: *Rolling Window Forecast Comparisons: Imports*

Horizon	1	2	3	4	5	6	7	8	9	10	11	12
TVP			1.002			0.980			0.932***			0.788***
Indicator	1.118	1.438***	2.389***	2.247***	2.999***	3.986***	4.110***	4.694***	3.611***	2.824***	2.799***	4.056***
GDPNow			0.530***									
Median	0.963	1.007	1.059*	1.081*	1.121***	1.091***	1.167***	1.053***	1.011***	1.034***	1.006*	1.019*
MA3	1.054	1.381	1.543***	1.527***	1.707***	1.621***	1.589***	1.489***	1.317***	1.327***	1.257***	1.206***
Hist. Avg.	0.974	1.020	1.072**	1.046	1.059*	1.049	1.068*	1.064**	1.020	1.022	1.011	0.965*
NoChange	1.513**	1.315	1.371***	1.415***	1.471***	1.534***	1.669***	1.742***	1.720***	1.766***	1.779***	1.734***

Notes*: MSE ratios* `MSE(model)/MSE(trimmed)` *across forecasting horizons in months based on four year rolling windows. See the text for details on the models. A ratio greater than one means that the forecast based on trimmed growth is more accurate. The trimmed growth rate omits components below the 65th percentile and above the 94th percentile of the disaggregated import and export growth distribution. Stars indicate statistical significance based on the Giacomini-White test at * $p < 0.1$, **$p < 0.05$,***$p < 0.01$. The null hypothesis is equal forecasting accuracy. Sample: 2002 M1 - 2024 M11.*

Table B6: *Rolling Window Forecast Comparisons: Exports*

Horizon	1	2	3	4	5	6	7	8	9	10	11	12
TVP			1.034			1.053			1.034			0.846***
Indicator	1.883***	2.611**	4.108***	3.221***	2.829***	3.779***	3.949***	3.701***	3.141***	2.809***	2.403***	3.572***
GDPNow			0.410**									
Median	2.023***	1.095**	1.239**	1.070***	1.085**	1.179***	1.169***	1.153**	1.226***	1.080***	1.034	1.135**
MA3	1.924***	2.882	2.645**	1.927***	1.719***	1.907***	1.850***	1.680***	1.535***	1.427***	1.302***	1.193***
Hist. Avg.	1.039	1.054***	1.036	0.994	1.004	1.055	1.077***	1.079***	1.100***	1.095***	1.074***	1.066***
NoChange	1.331	1.632	1.642***	1.553***	1.599***	1.725***	1.810***	1.855***	1.863***	1.867***	1.848***	1.837***

Notes*: MSE ratios* `MSE(model)/MSE(trimmed)` *across forecasting horizons in months based on four year rolling windows. See the text for details on the models. A ratio greater than one means that the forecast based on trimmed growth is more accurate. The trimmed growth rate omits components below the 65th percentile and above the 94th percentile of the disaggregated import and export growth distribution. Stars indicate statistical significance based on the Giacomini-White test at * $p < 0.1$, **$p < 0.05$,***$p < 0.01$. The null hypothesis is equal forecasting accuracy. Sample: 2002 M1 - 2024 M11.*

References

George Athanasopoulos, Puwasala Gamakumara, Anastasios Panagiotelis, Rob J Hyndman, and Mohamed Affan. Hierarchical forecasting. In *Macroeconomic Forecasting in the Era of Big Data: Theory and Practice*, pages 689–719. Springer, 2019.

Michael F Bryan and Stephen G Cecchetti. Measuring Core Inflation. In *Monetary policy*, pages 195–219. The University of Chicago Press, 1994.

Thomas R. Cook, Amaze Lusompa, and Johannes Matschke. Testing Hybrid Forecasts for Imports and Exports. *Economic Review*, June 2024.

David O Cushman. US Bilateral Trade Equations: Forecasts and Structural Stability. *Applied Economics*, 22(8):1093–1102, 1990.

Antonello D'Agostino, Michele Modugno, and Chiara Osbat. A Global Trade Model for the Euro Area. *International Journal of Central Banking*, 13(4):1–34, December 2017.

Francis Diebold and Roberto Mariano. Comparing Predictive Accuracy. *Journal of Business Economic Statistics*, 13(3):253–63, 1995.

Jim Dolmas et al. Trimmed Mean PCE Inflation. *Federal Reserve Bank of Dallas Working Paper*, 506, 2005.

Douglas Dunn, William Williams, and T.L. DeChaine. Aggregate Versus Subaggregate Models in Local Area Forecasting. *Journal of the American Statistical Association*, 71 (353):68–71, 1976.

Jon Faust and Jonathan H. Wright. Comparing Greenbook and Reduced Form Forecasts Using a Large Realtime Dataset. *Journal of Business & Economic Statistics*, 27(4): 468–479, None 2009.

Raffaella Giacomini and Halbert White. Tests of Conditional Predictive Ability. *Econometrica*, 74(6):1545–1578, 2006.

Roberto Golinelli and Giuseppe Parigi. Tracking World Trade and GDP in Real Time. *International Journal of Forecasting*, 30(4):847–862, 2014.

Christian Grimme, Robert Lehmann, and Marvin Noeller. Forecasting Imports with Information from Abroad. *Economic Modelling*, 98(C):109–117, None 2021.

Sandra Hanslin Grossmann and Rolf Scheufele. PMIs: Reliable Indicators for Exports? *Review of International Economics*, 27(2):711–734, May 2019.

David F Hendry and Kirstin Hubrich. Combining Disaggregate Forecasts or Combining Disaggregate Information to Forecast an Aggregate. *Journal of business & economic statistics*, 29(2):216–227, 2011.

Patrick C. Higgins. GDPNow: A Model for GDP 'Nowcasting' . FRB Atlanta Working Paper 2014-7, Federal Reserve Bank of Atlanta, Jul 2014.

Robert V Hogg. More light on the kurtosis and related statistics. *Journal of the American Statistical Association*, 67(338):422–424, 1972.

Gary Koop and Dimitris Korobilis. Bayesian Dynamic Variable Selection in High Dimensions. *International Economic Review*, 64(3):1047–1074, August 2023.

Robert Lehmann. Forecasting Exports Across Europe: What are the Superior Survey Indicators? *Empirical Economics*, 60(5):2429–2453, May 2021.

Òscar Jordà. Estimation and Inference of Impulse Responses by Local Projections. *American Economic Review*, 95(1):161–182, March 2005.