Working Papers of the Priority Programme 1859
Experience and Expectation.
Historical Foundations of Economic Behaviour
Edited by Alexander Nützenadel und Jochen Streb

No 26 (2020, October)

Behrens, Christoph

German Trade Forecasts since 1970 -
An Evaluation Using the Panel Dimension

Arbeitspapiere des Schwerpunktprogramms 1859 der Deutschen Forschungsgemeinschaft
„Erfahrung und Erwartung. Historische Grundlagen ökonomischen Handelns“ /
Working Papers of the German Research Foundation's Priority Programme 1859
"Experience and Expectation. Historical Foundations of Economic Behaviour"

HUMBOLDT-UNIVERSITÄT ZU BERLIN

Published in co-operation with the documentation and publication service of the Humboldt University, Berlin (https://edoc.hu-berlin.de).

ISSN: 2510-053X

Redaktion: Alexander Nützenadel, Jochen Streb, Ingo Köhler

V.i.S.d.P.: Alexander Nützenadel, Jochen Streb

SPP 1859 "Erfahrung und Erwartung. Historische Grundlagen ökonomischen Handelns"
Sitz der Geschäftsführung:
Humboldt-Universität
Friedrichstr. 191-193, 10117 Berlin
Tel: 0049-30-2093-70615, Fax: 0049-30-2093-70644
Web: https://www.experience-expectation.de
Koordinatoren: Alexander Nützenadel, Jochen Streb
Assistent der Koordinatoren: Ingo Köhler

Recommended citation:

Behrens, Christoph (2020): *German Trade Forecasts since 1970 - An Evaluation Using the Panel Dimension.* Working Papers of the Priority Programme 1859 "Experience and Expectation. Historical Foundations of Economic Behaviour" No 26 (October), Berlin

German Trade Forecasts since 1970
-
An Evaluation Using the Panel Dimension

Christoph Behrens[a,*]

October 2020

Abstract

I evaluate German export growth and import growth forecasts published by eight professional forecasters for the years 1971 to 2019. The focus of the evaluation is on the weak and strong efficiency as well as the unbiasedness of the forecasts. To this end, I use a novel panel-data set and estimate fixed-effects models taking into account panel-corrected standard errors. For the full time period, I find that both export and import growth forecasts are weakly but not strongly efficient. Unbiasedness depends on the forecast horizon being analyzed, with longer-term four-quarter-ahead forecasts being biased. I, furthermore, check for a possible change in forecasting behavior after incisive economic events in recent German history. I find that the strong efficiency of the forecasts did not change substantially over time. However, there is a change in forecasting behavior regarding the weak form of efficiency after the financial crisis 2008/2009.

JEL classification: C53; F17; F47

Keywords: Trade forecasts; German economic research institutes; Forecast Efficiency; Panel Data

Address:

[a] Department of Economics, Helmut Schmidt University, Holstenhofweg 85, P.O.B. 700822, 22008 Hamburg, Germany

[*] Corresponding author. E-mail: behrens.christoph@gmail.com

Acknowledgement:
I thank Christian Pierdzioch for very helpful and constructive feedback. I thank the German Research Foundation for financial support (Exploring the experience expectation nexus in macroeconomic forecasting using computational text analysis and machine learning; project number: 275693836).

1 Introduction

Economic recessions greatly influence international trade flows and trade policy. World trade is typically more severely affected by economic crises than GDP growth (see, among others, Freund, 2009; Levchenko et al., 2010; Berman and Martin, 2012; Chor and Manova, 2012, for documentations of the decline in world trade after the financial crisis in 2008/2009). One reason is, that often protectionist measures gain in popularity after strong economic recessions (Evenett, 2009; Kee et al., 2013). Durusoy et al. (2015), for instance, find that the number of export and import quotas and tariffs in the EU have substantially increased after 2008. Other authors find that the strong decline in trade after economic crises is caused by trade frictions (Behrens et al., 2013; Eaton et al., 2016) or the disruption of international production chains (Bems et al., 2010). As one of the world's largest exporters, Germany heavily relies on international trade. It is, therefore, crucial for policy and investment decisions that professional forecasters publish reliable, i.e., efficient and unbiased (Mincer and Zarnowitz, 1969; Holden and Peel, 1990), trade forecasts for Germany. I, therefore, analyze the efficiency and unbiasedness of German export and import growth forecasts for the years 1971 to 2019, published by eight professional forecasters.

The evaluation of macroeconomic forecasts for Germany has been the topic of numerous studies focusing on forecaster rankings (Sinclair et al., 2016), forecast revisions (Kirchgässner and Müller, 2006), the underlying assumptions of forecasts (Engelke et al., 2019), the economic value of forecasts (Döpke et al., 2018), forecast accuracy (Heilemann and Stekler, 2013), or forecast efficiency (Behrens et al., 2018a, 2020). The vast majority of the studies analyze GDP growth and inflation forecasts. The literature on the evaluation of trade forecasts, in contrast, is scarce, despite the importance of international trade for the German economy. Behrens (2019, 2020) analyzes German trade forecasts by means of machine learning techniques and finds evidence against the efficiency of some German forecasters. Rather than evaluating forecasters independently, I pool the data over all eight forecasters and focus on analyzing overall export growth and import growth forecasts for Germany over time, since 1970.

To this end, I build on earlier literature on the change of forecasting behavior over time. This literature has especially focused on forecasting behavior before and after the financial crisis of 2008/2009. This literature has considered different causes of forecasting-behavior changes, such as changes in the individuals responsible for the forecasts (Capistrán, 2008), changes in the loss function (Wang and Lee, 2014), or changes in the expectation-formation process of the forecasters (Frenkel et al., 2011; Pain et al., 2014). By means of a survey of German professional forecasters, Döpke et al. (2019b) find that forecasters tend to form more conservative forecasts after the financial crisis of 2008/2009. In a quantitative approach Döpke et al. (2019a) find only small differences in the forecasting behavior of German professional forecasters after the financial crisis. Again, the vast majority of the relevant literature analyzes GDP and inflation forecasts. In order to check for a possible change in forecasting behavior regarding trade forecasts, I evaluate subsamples after incisive economic events in recent German history, namely the oil price shocks in the early and late 1970s, German reunification, and the financial crisis of 2008/2009.

Keane and Runkle (1990) argue in an early application of panel-regressions to forecast evaluation that pooling forecasts results in a more efficient evaluation of forecast rationality. Hence, I build on research by Döpke and Fritsche (2006) and Döpke et al. (2019a), who analyze German GDP and inflation forecasts by means of fixed-effects-panel regressions. To this end, I follow Döpke et al. (2019a) and use Beck and Katz (1995) panel-corrected standard errors (PCSE), which have better finite sample properties for time-series cross-sectional data than the also common feasible generalized least squares (FGLS) estimator developed by Parks (1967).

I structure the remainder of the paper as follows: I present the data in Section 2. The empirical analysis in Section 3 consists of a brief description of the estimation technique and tests for efficiency as well as unbiasedness of the forecasts for the full sample and for subsamples corresponding to incisive economic events in recent German history. In Section 4 I conclude.

2 Data

I use a modified version of a novel data set, which has recently been analyzed in nonparametric forecast-evaluation studies by Behrens (2019, 2020). It consists of annual export growth and import growth forecasts for the years 1971 to 2019 published by seven German economic research institutes and one collaboration of economic research institutes. Five of the forecasters are among the largest politically and economically independent German economic research institutes, namely (in alphabetical order): Deutsches Institut für Wirtschaftsforschung Berlin (DIW), Hamburgisches Weltwirtschaftsarchiv/-institut (HWWI)[1], ifo Institut für Wirtschaftsforschung Munich (ifo), Institut für Weltwirtschaft Kiel (IfW), and Rheinisch-Westfälisches Institut für Wirtschaftsforschung Essen (RWI).[2] Two of the forecasters receive funding from interest groups, i.e., Institut für Makroökonomie und Konjunkturforschung Düsseldorf (IMK), which is financed by labor unions, and Institut der deutschen Wirtschaft Köln (IW), which is financed by employer's associations. In addition, the list of forecasters comprises a collaboration of the leading economic research institutes in Germany, the so called joint forecast or Gemeinschaftsdiagnose (GD).

The research institutes publish forecasts midyear and at the turn of a year. The former forecasts have a forecast horizon of two quarters and predict trade aggregates for the respective current year, whereas the latter have a forecast horizon of four quarters and predict trade aggregates for the respective upcoming year. The the total number of forecasts as well as the exact publication dates vary across forecasters, resulting in a possible information advantage of forecasters who publish their forecasts at later dates. I follow Döpke and Fritsche (2006) and Döpke et al. (2019a) and account for this issue by means of a fixed-effects-panel regression (see Section 3.1). In order to compute forecast errors, I use realized values of German export and import growth, as published by the German statistical office.[3] I use initial release national accounts data to minimize the effects of data revisions. Furthermore, I adjust the reference time-series for every forecaster, as the economic research institutes switch from forecasts for West-Germany to forecasts for reunified Germany at different points in time between 1992 and 1993. Following Behrens et al.

[1] HWWI became a privately funded institute in 2006.

[2] The sixth main German economic research institute (Institut für Wirtschaftsforschung Halle) is omitted, because it has only been publishing forecasts since German reunification.

[3] Data taken from "Wirtschaft und Statistik" publications: https://www.destatis.de/EN/Publications/WirtschaftStatistik/WirtschaftStatistik.html

Table 1: Descriptive Statistics of Trade Forecast Errors

Institute	Forecast	N	ME	RMAE	RMSE	$e<0$	N	ME	RMAE	RMSE	$e<0$
		Two-Quarters-Ahead					*Four-Quarters-Ahead*				
DIW	*Exports*	38	0.218	1.348	2.272	0.45	49	-0.631	1.977	5.027	0.59
HWWI		33	-0.091	1.321	2.369	0.58	44	0.216	1.910	4.721	0.48
ifo		43	0.107	1.314	2.222	0.49	45	-0.349	1.891	4.653	0.58
IfW		42	0.188	1.344	2.437	0.40	47	-0.740	1.818	4.135	0.55
RWI		23	0.598	1.250	2.681	0.48	24	0.021	1.728	3.731	0.50
IMK		12	-0.592	1.228	2.082	0.50	42	-0.369	1.955	4.821	0.55
IW		19	-0.139	1.491	2.970	0.47	36	0.043	1.912	4.981	0.53
GD		48	-0.448	1.708	3.727	0.54	49	-0.017	2.001	5.265	0.53
Pooled		258	0.001	1.410	2.710	0.49	336	-0.257	1.912	4.743	0.54
		Two-Quarters-Ahead					*Four-Quarters-Ahead*				
DIW	*Imports*	38	0.650	1.410	2.824	0.42	49	-0.094	1.737	3.863	0.53
HWWI		33	0.236	1.421	2.696	0.39	43	0.509	1.765	3.946	0.42
ifo		43	0.021	1.264	2.214	0.42	45	0.036	1.632	3.384	0.49
IfW		42	0.521	1.358	2.319	0.38	47	0.294	1.663	3.399	0.49
RWI		23	0.867	1.447	3.400	0.39	24	0.592	1.676	3.955	0.50
IMK		12	-0.108	1.255	2.179	0.50	42	0.005	1.692	3.623	0.50
IW		19	0.274	1.467	2.711	0.42	28	-0.159	1.713	3.360	0.50
GD		49	0.076	1.527	3.101	0.47	47	0.242	1.788	4.186	0.49
Pooled		259	0.320	1.404	2.718	0.42	325	0.166	1.711	3.782	0.49

Notes: N: Number of observations. Mean error: ME = $\frac{1}{T}\sum_{t=1}^{T} e_t$. Root mean absolute error: RMAE = $\sqrt{\frac{1}{T}\sum_{t=1}^{T}|e_t|}$. Root mean squared error: RMSE = $\sqrt{\frac{1}{T}\sum_{t=1}^{T}{e_t}^2}$. $e<0$: Share of negative forecast errors.

(2018b), I compute forecast errors by subtracting the realized values for German export or import growth from the forecasted value of a given year, such that:

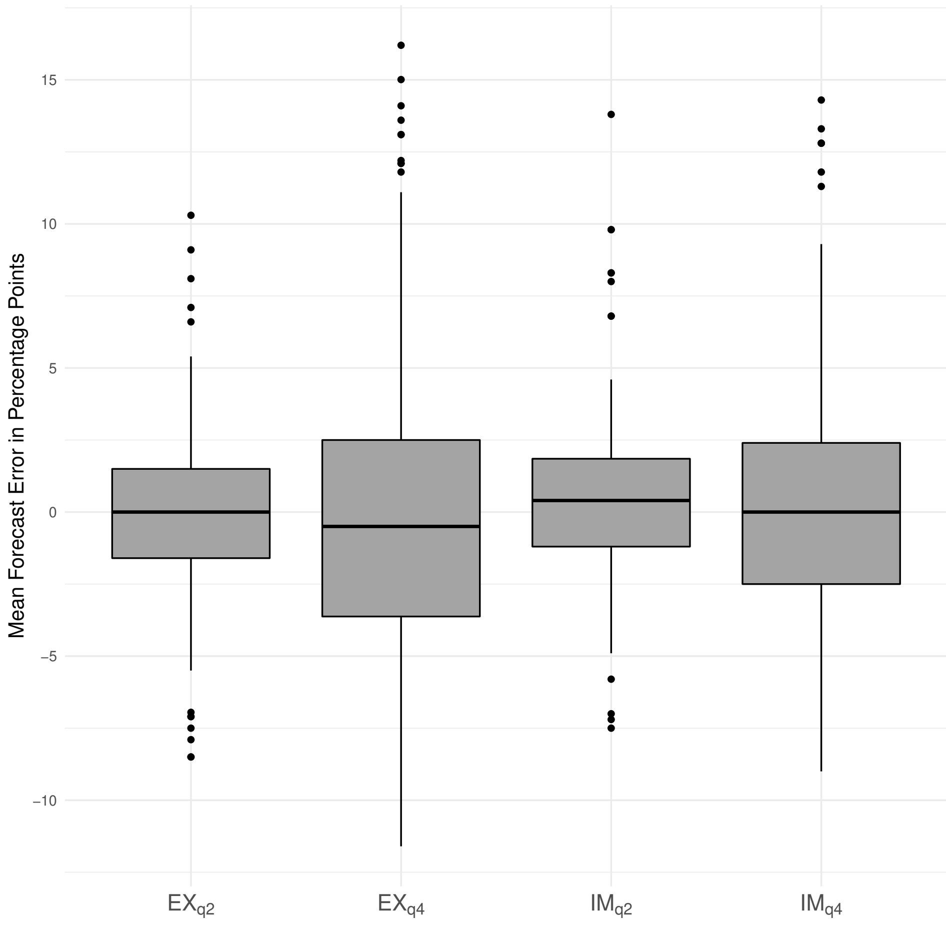
$$e_{i,t(h)} = \hat{y}_{i,t(h)} - y_t. \tag{1}$$

Here, $e_{i,t(h)}$ denotes the forecast error made by economic research institute i for the year $t = 1971-2019$ at a forecast horizon of $h = 2,4$ quarters and $\hat{y}_{i,t(h)}$ denotes the export growth or import growth forecast published by institute i in year t, which also depends on the forecast horizon, h. Finally, y_t denotes the realized value of German export or import growth for year t for which the forecast was formed.

Table 1 reports descriptive statistics of the analyzed forecast errors. There are more observations available for the longer-term forecasts than for the shorter two-quarter-ahead forecasts. IMK contributes the fewest observations, namely 12, for both trade aggregates at the two quarter forecast horizon, whereas most observations are available for DIW, namely 49 for the longer-term forecast horizon for both trade aggregates. The pooled datasets contain 258 (259) observations for the two-quarters-ahead export (import) growth forecasts and 336 (325) for the four-quarters-ahead export (import) growth forecasts. As is to be expected, the root mean absolute error (RMAE) and the root mean squared error (RMSE) statistics are higher for the longer-term forecast horizon. Furthermore, all error statistics are generally higher than the values observed in the more common studies of GDP and inflation forecasts (see, among others, Döpke et al., 2017; Behrens et al., 2018a). This is due to the fact that trade aggregates are among the most volatile components of German national accounts statistics and are, therefore, harder to predict (Döhrn and Schmidt, 2011). The share of negative forecast errors should equal 0.5 if the forecasters, on average, do not overestimate or underestimate. For the four-quarter-ahead export growth forecasts and the two-quarters-ahead import growth forecasts, the value deviates the most (i.e., by 0.06 and 0.08 percentage points) from the 0.50 value, hinting at a possible bias. This can also be observed in Figure 1, which depicts boxplots of the data. The boxplots depict the distribution of export and import growth forecast errors at both forecast horizons. For the two-quarters-ahead-forecasts, denoted by EX_{q2} and IM_{q2}, the forecast errors are more closely distributed around the mean, resulting in more narrow boxplots. A longer forecast horizon causes a larger deviation from the mean and hence broader boxplots. Furthermore, the boxplot for the four-quarters-ahead export growth forecasts, EX_{q4}, is shifted for a larger part below zero, again hinting at a possible underestimation bias.

In order to model the information set of the economic research institutes at forecast formation, I use typical trade variables as well as other macroeconomic variables commonly used to predict economic growth. All variables enter the list of predictors in normalized form. In doing so, I build on research by D'Agostino et al. (2017), who show that incorporating both types of macroeconomic aggregates improve trade forecasts for the euro area. In order to minimize the effects of data revisions, I use a backward-looking moving average of order 12 (see also, Behrens et al., 2020). Based on a study by Drechsel and Scheufele (2012), I, furthermore, take publication lags of the variables into account. In general, I assume that when a forecast is published, for

Figure 1: Boxplots of Pooled Trade Forecast Errors



Note: EX_{q2}: Two-quarter-ahead export growth forecast. EX_{q4}: Four-quarter-ahead export growth forecast. IM_{q2}: Two-quarter-ahead import growth forecast. IM_{q4}: Four-quarter-ahead import growth forecast.

instance, in January, it is based on information available in December (Behrens et al., 2018b). The list of predictors to proxy the forecasters' information set includes:

- *Industrial Orders*: The year-on-year rate of change of the industrial orders received for Germany indicates demand fluctuations (see, among others, Döpke et al., 2017, on using industrial orders, inter alia, to predict German recessions).

- *Unemployment rate*: The monthly German unemployment rate in percent of civilian labor is included following Behrens (2020) who finds evidence, using nonparamteric techniques, that the unemployment rate might not be efficiently incorporated in German trade forecasts (see also, D'Agostino et al., 2017, on improving trade forecasts by means of macroeconomic variables such as the unemployment rate).

- *Business climate*: The monthly ifo tendency survey for German manufacturing enters the list of predictors. Studies by Frale et al. (2010) and Lehmann (2015) suggest that survey data is essential for the forecasting of exports in Europe.

- *Production Germany*: Year-on-year rate of change of the monthly German total manufacturing output. I follow Behrens et al. (2018a,b) who evaluate the efficiency of German GDP growth and inflation forecasts by means of machine learning techniques.

- *Production G7*: The year-on-year rate of change of the monthly total manufacturing output of the G7 is added as a leading indicator of the economic development of Germany's main trading partners, which is a crucial information for forecasters as Campos et al. (2019) show that international business cycles are oftentimes synchronized (see also, Guichard and Rusticelli, 2011, on improving trade forecasts by means of industrial production indices).

- *Oil price*: Year-on-year rate of change of the monthly crude oil price (WTI) in dollars per barrel. I follow Döpke et al. (2019a) in using the oil price as a proxy for input prices.

- *Leading Indicator*: The monthly OECD composite leading indicator for Germany enters the set of predictors. Heinisch and Scheufele (2018) use the OECD leading indicator for Germany, inter alia, to forecast the German GDP.

- *Real effective exchange rate (REER)*: Year-on-year rate of change of the monthly narrow effective exchange rate for Germany (CPI-based). The REER serves as a measure of the international price competitiveness of Germany (Grimme et al., 2018; Lehmann, 2015).

- *Trade Policy Uncertainty Index (TPU)*: Monthly measure of media[4] attention to news related to trade policy uncertainty. The TPU, developed by Caldara et al. (2019), is included as a measure of uncertainty regarding international trade.

3 Empirical Analysis

3.1 Forecast Efficiency and Unbiasedness Tests

In order to test for weak and strong efficiency as well as unbiasedness of German export and import growth forecasts, I follow Döpke et al. (2019a), who build on research by Keane and Runkle (1990) as well as Döpke and Fritsche (2006), and implement the Holden and Peel (1990) approach to testing forecast efficiency and unbiasedness by means of a fixed-effects panel-regression. Holden and Peel (1990) define a strong and weak form of efficiency, where the former holds if the forecast error cannot be explained by information available to a forecaster at the time of forecast formation. The latter form of efficiency holds if a forecast error cannot be explained by its preceding forecast error (see also, Öller and Barot, 2000; Timmermann, 2007; Behrens et al., 2020). I implement tests for weak and strong efficiency as well as unbiasedness of the export and import growth forecasts by means of the following regression model:

$$e_{i,t(h)} = \beta_0 + \beta_1 e_{i,t(h)-1} + \beta_j \mathbf{X}_{j,i,t(h)-h} + \alpha_i + \lambda_{t(h)} + u_{i,t(h)}. \tag{2}$$

[4]For the construction of the trade policy uncertainty index, electronic archives of 7 leading U.S. newspapers discussing trade policy uncertainty, namely Boston Globe, Chicago Tribune, Guardian, Los Angeles Times, New York Times, Wall Street Journal, and Washington Post, are analyzed by means of automated text-search (for details on the construction of the index, see Caldara et al., 2019).

Here, $e_{i,t(h)}$ again denotes the forecast error made by economic research institute i for the year $t = 1971 - 2019$ with forecast horizon $h = 2,4$ quarters. $\mathbf{X}_{j,i,t(h)-h}$ is the vector of j predictors, available to institute i in period $t(h) - h$, when the forecast was formed, which depends on the forecast horizon. $\lambda_{t(h)}$ and α_i are time and entity fixed effects. $e_{i,t(h)-1}$ is the error of the forecast of the previous year and $u_{i,t(h)}$ is the statistical error term.[5]

As in Holden and Peel (1990), a forecast is considered as unbiased if the coefficient of the intercept is not statistically significantly different from zero, i.e., if the null hypothesis, $H_0 : \beta_0 = 0$, cannot be rejected. Strong efficiency of forecasts holds if the predictors do not have statistically significant explanatory power for the forecast error, i.e., if the null hypotheses $H_0 : \beta_2 = 0, \beta_3 = 0, ..., \beta_j = 0$ cannot be rejected. Analogously, a forecast is weakly efficient if the null hypothesis $H_0 : \beta_1 = 0$ cannot be rejected, i.e., if the lagged forecast error is uncorrelated to the forecast error.

I use both time and entity fixed effects in the regression model. In doing so, I control for effects that equally affect all entities (i.e., economic research institutes) but change over time, such as oil price shocks in the 1970s, as well as effects that are stable over time but change across entities, such different forecast models or economic theories of the institutes. The former time fixed effects, $\lambda_{t(h)}$, can be interpreted as the element of surprise of a given year, which should have strong influence on the forecast error when a crisis hits the economy for the first time. The latter entity fixed effects, α_i, in contrast, control for slightly differing forecast horizons due to different publication dates of the economic research institutes (see also Döpke and Fritsche, 2006; Döpke et al., 2019a).

Keane and Runkle (1990) argue that, when one analyzes the efficiency of forecasts by means of a panel-regression model, the model needs to account for heteroskedasticity. A common way to address this issue is to use the feasible generalized least squares estimator (Parks, 1967; Kmenta, 1986), which is sometimes referred to as the Parks estimator. However, Beck and Katz (1995)

[5] I roughly follow the notation used by Stock and Watson (2007).

introduced the so called panel-corrected standard-errors estimator and showed that it performs significantly better than the Parks estimator for finite samples. The PCSE estimator takes into account non-spherical errors, i.e., it is robust to unit heteroskedasticity as well as contemporaneous correlation across units. The latter characteristic is especially crucial for this study, as it is reasonable to assume that forecast errors are correlated across economic research institutes. The PCSE estimator is similar to other heteroskedasticity consistent estimators, such as the ones brought forward by Huber (1967), White (1980) or MacKinnon and White (1985), yet it is specifically designed for time-series-cross-section (TSCS) data, having more observations along the time-series dimension than the cross-section dimension (i.e., $T > N$). I, therefore, follow Döpke et al. (2019a) and implement the PCSE estimator by Beck and Katz (1995) in my empirical analysis. Due to the $T > N$-nature of the data, it is not necessary to control for a possible Nickell-bias (Nickell, 1981; Arellano and Bond, 1991), even though the lagged dependent variable is included in the estimation equation (Eq. 2), as was shown by Gaibulloev et al. (2014).

I use the R programming environment for statistical computing (R Core Team, 2020) to estimate the fixed effects model, and I use the add-on package "pcse" (Bailey and Katz, 2011) to compute Beck and Katz (1995) panel-corrected standard errors. Tables 2 and 3 present results of efficiency tests for the full sample of export and import growth forecasts, respectively. Regarding two-quarter-ahead export growth forecasts (Table 2, top panel), I find evidence against the strong form of efficiency. The predictors business climate, OECD leading indicator, and real effective exchange rate have statistically significant explanatory power for the forecast error. The forecasts are unbiased as, for all but one specification, the coefficients of the intercept terms are insignificant. Furthermore, the coefficient of the lagged forecast error is not significant in any specification, indicating that the two-quarters-ahead export growth forecasts are weakly efficient. Tables 2 and 3 also show adjusted R^2-statistics for the analyzed regression models (Eq. 2) with time and entity fixed effects as well as for regression models using only entity fixed effects.[6] For two-quarter-ahead export growth forecasts, the adjusted R^2-statistics of the standard specifi-

[6] Detailed results of the regression models using only entity fixed effects are not reported but are available from the author upon request.

Table 2: Efficiency Tests of Export Growth Forecasts for Germany - Full Sample

	Dependent Variable: Export Growth Forecast Error (Two-Quarters-Ahead)									
Intercept	2.255 (1.512)	2.011 (7.605)	1.617 (1.371)	1.054 (1.495)	1.388 (1.419)	1.356 (1.424)	1.752 (1.255)	3.962** (1.668)	1.422 (1.431)	1.394 (1.417)
Industr. Orders	0.955 (0.665)									
Unemp. Rate		0.328 (3.979)								
Busin. Climate			0.571** (0.243)							
Production GER				0.093 (0.156)						
Production G7					0.684 (0.781)					
Oil Price						-0.228 (0.446)				
Leading Indic.							2.125*** (0.760)			
REER								-1.731** (0.686)		
TPU									0.043 (0.276)	
e_{t-1}	-0.049 (0.065)	-0.057 (0.065)	-0.050 (0.065)	-0.056 (0.065)	-0.051 (0.066)	-0.059 (0.066)	-0.057 (0.063)	-0.024 (0.064)	-0.055 (0.066)	-0.059 (0.066)
Observations	258	258	258	258	258	258	258	258	258	258
Adjusted R^2	0.784	0.780	0.786	0.781	0.781	0.780	0.792	0.789	0.780	0.781
Adjusted R^2 **(w/o time FE)**	0.001	-0.017	-0.014	0.015	0.037	-0.005	-0.017	0.049	0.005	-0.013
	Dependent Variable: Export Growth Forecast Error (Four-Quarters-Ahead)									
Intercept	-4.157 (0.438)	9.514** (4.433)	-0.950** (0.439)	-2.365*** (0.811)	-1.489*** (0.491)	-0.774** (0.429)	-0.910** (0.449)	-0.627 (1.026)	-0.930** (0.450)	
Industr. Orders	2.116*** (0.480)									
Unemp. Rate		5.566 (2.347)								
Busin. Climate			0.800*** (0.191)							
Production GER				1.142** (0.553)						
Production G7					1.206** (0.525)					
Oil Price						0.725*** (0.193)				
Leading Indic.							1.898*** (0.287)			
REER								-0.145 (0.409)		
TPU									0.058 (0.270)	
e_{t-1}	0.000 (0.041)	0.006 (0.041)	-0.003 (0.041)	-0.013 (0.041)	-0.009 (0.041)	0.009 (0.040)	0.022 (0.039)	-0.005 (0.042)	-0.006 (0.042)	-0.006 (0.042)
Observations	336	336	336	336	336	336	336	336	336	336
Adjusted R^2	0.944	0.941	0.944	0.941	0.941	0.944	0.948	0.940	0.940	0.940
Adjusted R^2 **(w/o time FE)**	0.003	0.001	0.115	0.021	0.000	0.014	0.035	0.002	0.007	0.003

Notes: Results are computed by means of an entity and time fixed effects regression with panel-corrected standard errors (PCSE, Beck and Katz, 1995). PCSE in parentheses. Entity and time fixed effects are excluded to save journal space. W/o time FE: Model specification using only entity fixed effects. ***, **, * denote statistical significance at the 1, 5, 10 %-level.

cation vary around approximately 0.78, whereas the adjusted R^2-statistics of the model without time fixed effects range between -0.017 and 0.046. This indicates that a large part of the explanatory power of the forecast error is linked to time fixed effects, i.e. the year for which a forecast was formed. The time fixed effects can be interpreted as the element of surprise of a given year. The results show the importance of using a time and entity fixed regression model, when analyzing trade forecast errors.

The same holds for the four-quarters-ahead export growth forecasts, for which the adjusted R^2-statistics of the standard specification vary around approximately 0.94, whereas the adjusted R^2-statistics of the model without time fixed effects range between 0 and 0.115. Furthermore, these forecasts are weakly efficient since the lagged forecast errors do not have significant explanatory power for the forecast error. However, for the four-quarters-ahead forecast horizon, the forecasts are biased, as was already indicated by the boxplot shown in Figure 1. The coefficients of the intercepts are statistically significant for all specifications except for the regressions using the real effective exchange rate and industrial orders. There is also evidence against the strong form of forecast efficiency, as the coefficients of several predictors have significant explanatory power for the forecast error, namely the coefficients of the predictors industrial orders, business climate, German and G7 production, oil price, and OECD leading indicator.

Next, I turn to two- and four-quarters-ahead import growth forecasts for Germany, which are reported in the top and bottom panel of Table 3. The forecasts with a shorter forecast horizon are unbiased and weakly efficient, as neither the coefficients of the intercepts nor of the lagged forecast errors are statistically significant. Yet, I reject the strong form of efficiency, as the shorter-term import growth forecast error is linked to industrial orders, business climate, G7 production, the OECD leading indicator, and the real effective exchange rate. When comparing the adjusted R^2-statistics of the standard specification and the adjusted R^2-statistics of the model without time fixed effects, I again find that forecast errors are explained to a large part by time fixed effects, as the former R^2-statistics vary around 0.79, whereas the latter range between -0.02 and 0.046.

Table 3: Efficiency Tests of Import Growth Forecasts for Germany - Full Sample

	Dependent Variable: Import Growth Forecast Error (Two-Quarters-Ahead)									
Intercept	1.173 (1.300)	-2.284 (6.901)	0.171 (1.213)	-0.551 (1.303)	0.007 (1.225)	0.003 (1.255)	0.465 (1.090)	1.873 (1.529)	0.071 (1.266)	0.012 (1.248)
Industr. Orders	1.265** (0.586)									
Unemp. Rate		-1.221 (3.601)								
Busin. Climate			0.376* (0.226)							
Production GER				0.794 (0.643)						
Production G7					1.153* (0.689)					
Oil Price						-0.247 (0.409)				
Leading Indic.							2.006*** (0.698)			
REER								-1.219* (0.625)		
TPU									0.086 (0.251)	
$\mathbf{e_{t-1}}$	-0.006 (0.070)	-0.014 (0.071)	-0.010 (0.070)	-0.016 (0.070)	-0.009 (0.070)	-0.015 (0.071)	0.022 (0.071)	0.004 (0.070)	-0.012 (0.071)	-0.013 (0.071)
Observations	259	259	259	259	259	259	259	259	259	259
Adjusted R^2	0.800	0.790	0.793	0.792	0.794	0.790	0.801	0.795	0.790	0.791
Adjusted R^2 (w/o time FE)	-0.018	-0.020	0.046	0.001	0.005	0.025	0.001	0.031	-0.014	-0.016
	Dependent Variable: Import Growth Forecast Error (Four-Quarters-Ahead)									
Intercept	-1.570*** (0.462)	-4.300 (4.113)	-2.077*** (0.462)	-3.219*** (0.738)	-2.439*** (0.494)	-1.990*** (0.463)	-2.032*** (0.467)	-0.717 (0.981)	-2.117*** (0.476)	-2.083*** (0.458)
Industr. Orders	1.768*** (0.449)									
Unemp. Rate		-1.175 (2.173)								
Busin. Climate			-0.152 (0.195)							
Production GER				0.942* (0.483)						
Production G7					0.853* (0.485)					
Oil Price						0.326 (0.199)				
Leading Indic.							1.521*** (0.312)			
REER								-0.076 (0.264)		
TPU									-0.118 (0.215)	
$\mathbf{e_{t-1}}$	-0.034 (0.046)	-0.051 (0.046)	-0.048 (0.046)	-0.045 (0.046)	-0.045 (0.046)	-0.044 (0.046)	-0.039 (0.044)	-0.041 (0.046)	-0.049 (0.046)	-0.049 (0.046)
Observations	325	325	325	325	325	325	325	325	325	325
Adjusted R^2	0.915	0.911	0.911	0.912	0.912	0.912	0.917	0.912	0.911	0.912
Adjusted R^2 (w/o time FE)	0.024	-0.006	0.159	0.059	0.004	0.030	0.034	-0.011	-0.011	-0.008

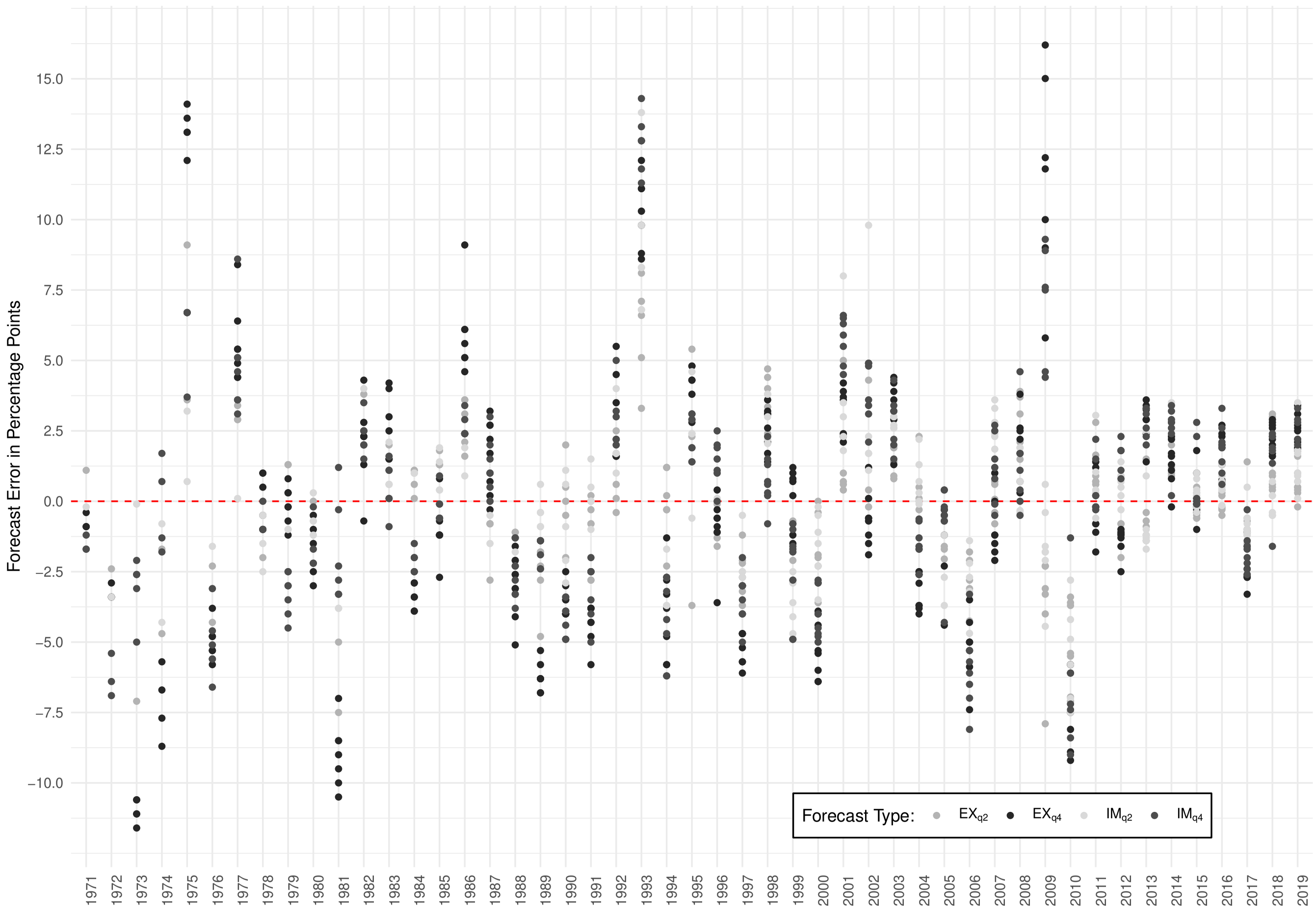
Notes: Results are computed by means of an entity and time fixed effects regression with panel-corrected standard errors (PCSE, Beck and Katz, 1995). PCSE in parentheses. Entity and time fixed effects are excluded to save journal space. W/o time FE: Model specification using only entity fixed effects. ***, **, * denote statistical significance at the 1, 5, 10 %-level.

The four-quarters-ahead import growth forecast errors, reported in the bottom panel of Table 3, are biased just as their export growth counterparts. There is strong statistical significance of the coefficients of the intercept terms in all but two specifications. I reject the strong form of efficiency due to statistically significant dependency of the forecast error on industrial orders, German and G7 production, as well as the OECD leading indicator. The weak form of efficiency cannot be rejected, because the lagged forecast errors do not have explanatory power for the forecast error. The forecast error is, however, explained to a large part by time fixed effects as is indicated by the large difference of the adjusted R^2-statistics of the regression models with and without time fixed effects.

3.2 Crises Subsamples

In the spirit of the literature on the change of forecasting behavior after economic recessions and especially the financial crisis of 2008 (Frenkel et al., 2011; Döpke et al., 2019a,b), on the one hand, and the literature on the severe effects of economic recessions on international trade and protectionist measures (Levchenko et al., 2010; Chor and Manova, 2012; Kee et al., 2013; Eaton et al., 2016), on the other hand, I define several subsamples of the dataset. In doing so, I capture possible changes in forecasting behavior in different periods of recent German history. In contrast to the time fixed effects in the estimation equation (Eq. 2), which capture the effects of a single year on the forecast errors, the sample split captures broader time trends.

Figure 2: Pooled Trade Forecast Errors over Time



Note: EX_{q2}: Two-quarter-ahead export growth forecast. EX_{q4}: Four-quarter-ahead export growth forecast. IM_{q2}: Two-quarter-ahead import growth forecast. IM_{q4}: Four-quarter-ahead import growth forecast.

I split the data based on Figure 2, which plots German export growth and import growth forecast errors over time. Forecast errors produced by longer-term forecasts are depicted in dark grey, whereas those errors produced by two-quarter-ahead forecasts are depicted in light grey. The first subsample comprises forecast errors for the years 1971 to 1981. It is the smallest of the analyzed subsamples, and it is characterized by relatively high and scattered forecast errors. Due to the two oil price shocks in that period, I will refer to it as the "*oil crises*" subsample. The next subsample spans from 1982 until 1992, and it is characterized by less widespread forecast errors than the time period before. This subsample will be referred to as "*West Germany*". The subsample after all forecasters switched from forecasts for West Germany to forecasts for reunified Germany until the financial crisis, i.e., 1994 to 2008, will be referred to as "*reunified Germany*". Finally, the last subsample, referred to as the "*post financial crisis*" subsample, contains forecast errors for the years 2010 to 2019 and is the shortest subsample. The forecast errors for the years 1993 and 2009 are omitted as all forecasters produce very large forecast errors of up to 15 percentage points, due to high uncertainty after German reunification and the financial crisis of 2008 (see also, Döpke et al., 2019a, for a similar approach regarding forecast errors after the financial crisis). All subsamples continue to be of the $T > N$-type, such that Equation (2) can still be estimated by means of PCSE by Beck and Katz (1995). Tables 4 and 5 report results of efficiency and unbiasedness tests for all subsamples for four-quarter-ahead export and import growth forecasts, respectively.

Beginning with the export growth forecasts, it can be assessed that the results of the subsamples do not differ severely from the ones computed by means of the full sample. For the "*oil crises*" subsample, the results are very similar to the full sample, as forecasts are biased but weakly efficient and there is evidence against the strong form of efficiency. Regarding the forecast errors for the years 1982 to 1992, there continues to be evidence against the unbiasedness of export growth forecasts. I reject the strong form of efficiency and, different from previous samples, I find some evidence against the weak form of efficiency. This finding also holds for the subsample "*reunified Germany*", for which some of the coefficients of the lagged forecast error are also statistically significant. Furthermore, there is less evidence against the unbiasedness of forecasts

Table 4: Efficiency Tests of Four-Quarters-Ahead Export Growth Forecasts for Germany - Subsamples

	Subsample: Oil Crises (1971 - 1981)									
Intercept	-0.791*	5.823	-0.943**	-0.989	-0.980*	-1.038**	-0.965**	-1.682*	-1.151**	-0.966**
	(0.403)	(7.574)	(0.350)	(1.154)	(0.485)	(0.461)	(0.383)	(0.987)	(0.454)	(0.374)
Industr. Orders	0.675									
	(0.960)									
Unemp. Rate		3.617								
		(4.036)								
Busin. Climate			0.674**							
			(0.316)							
Production GER				0.019						
				(0.860)						
Production G7					0.032					
					(0.629)					
Oil Price						-0.330				
						(1.217)				
Leading Indic.							0.949**			
							(0.379)			
REER								0.314		
								(0.404)		
TPU									-0.395	
									(0.498)	
$\mathbf{e_{t-1}}$	0.007	0.012	-0.014	0.008	0.008	0.009	0.010	0.007	0.005	0.008
	(0.035)	(0.034)	(0.036)	(0.035)	(0.035)	(0.035)	(0.034)	(0.035)	(0.036)	(0.035)
Observations	56	56	56	56	56	56	56	56	56	56
Adjusted R^2	0.985	0.985	0.986	0.985	0.985	0.985	0.987	0.985	0.985	0.985
Adjusted R^2 (w/o time FE)	-0.047	0.056	0.159	-0.071	-0.071	0.260	-0.052	-0.070	-0.071	-0.049
	Subsample: West Germany (1982 - 1992)									
Intercept	5.898***	6.523***	4.015**	7.547***	5.342**	4.121**	8.119***	4.516	4.008**	3.969**
	(1.810)	(2.340)	(1.669)	(2.009)	(2.047)	(1.637)	(2.412)	(3.628)	(1.650)	(1.647)
Industr. Orders	0.502**									
	(1.063)									
Unemp. Rate		5.763								
		(3.533)								
Busin. Climate			0.051							
			(0.304)							
Production GER				3.381***						
				(1.175)						
Production G7					1.417					
					(1.274)					
Oil Price						0.325				
						(0.517)				
Leading Indic.							3.882**			
							(1.604)			
REER								0.182		
								(0.979)		
TPU									0.246	
									(0.669)	
$\mathbf{e_{t-1}}$	0.280*	0.253	0.231	0.272*	0.277*	0.239	0.191	0.231	0.234	0.233
	(0.157)	(0.160)	(0.164)	(0.152)	(0.165)	(0.162)	(0.156)	(0.161)	(0.163)	(0.163)
Observations	79	79	79	79	79	79	79	79	79	79
Adjusted R^2	0.923	0.921	0.918	0.925	0.920	0.919	0.924	0.918	0.918	0.920
Adjusted R^2 (w/o time FE)	-0.001	0.050	0.342	-0.106	0.079	-0.010	0.169	0.260	-0.111	-0.095

Continued on next page.

Continued from previous page.

	Subsample: Reunified Germany (1994 - 2008)									
Intercept	0.197 (1.535)	-3.301 (2.406)	-2.537** (0.931)	0.814 (1.253)	-1.452 (1.084)	-2.735*** (1.010)	0.071 (1.220)	-1.062 (1.423)	-2.613** (1.055)	-2.644*** (0.957)
Industr. Orders	1.827** (0.761)									
Unemp. Rate		1.103 (3.846)								
Busin. Climate			0.608 (0.400)							
Production GER				1.523 (1.089)						
Production G7					1.647 (1.095)					
Oil Price						-0.154 (0.422)				
Leading Indic.							2.195*** (0.708)			
REER								-1.201* (0.716)		
TPU									-0.029 (0.589)	
e_{t-1}	-0.129 (0.085)	-0.156* (0.087)	-0.137 (0.086)	-0.155* (0.086)	-0.144* (0.086)	-0.166* (0.087)	-0.108 (0.085)	-0.133 (0.084)	-0.162* (0.088)	-0.162* (0.087)
Observations	111	111	111	111	111	111	111	111	111	111
Adjusted R^2	0.882	0.876	0.880	0.879	0.879	0.877	0.885	0.880	0.876	0.878
Adjusted R^2 (w/o time FE)	0.171	0.306	0.112	0.125	0.103	0.231	0.162	0.103	0.107	0.111
	Subsample: Post Financial Crisis (2010 - 2019)									
Intercept	-1.235 (2.515)	-3.930 (4.294)	-5.121*** (1.008)	-0.728 (2.120)	0.104 (2.706)	-5.191*** (1.312)	-4.837*** (1.012)	-4.934*** (1.093)	-4.912*** (1.065)	-4.933*** (1.080)
Industr. Orders	0.816 (0.560)									
Unemp. Rate		-2.231 (8.455)								
Busin. Climate			0.258 (0.466)							
Production GER				0.840** (0.338)						
Production G7					1.005** (0.495)					
Oil Price						-0.120 (0.242)				
Leading Indic.							0.451 (0.418)			
REER								-0.803 (0.895)		
TPU									0.128 (0.176)	
e_{t-1}	-0.226** (0.090)	-0.232** (0.113)	-0.202* (0.105)	-0.265*** (0.096)	-0.252*** (0.094)	-0.202* (0.107)	-0.182* (0.105)	-0.235* (0.096)	-0.219** (0.090)	-0.221** (0.091)
Observations	76	76	76	76	76	76	76	76	76	76
Adjusted R^2	0.933	0.931	0.931	0.935	0.934	0.931	0.932	0.932	0.932	0.932
Adjusted R^2 (w/o time FE)	0.328	0.656	0.311	0.349	0.508	0.339	0.300	0.362	0.358	0.311

Notes: Results are computed by means of an entity and time fixed effects regression with panel-corrected standard errors (PCSE, Beck and Katz, 1995). Dependent variable: Export growth forecast error (four-quarters-ahead). PCSE in parentheses. Entity and time fixed effects are excluded to save journal space. W/o time FE: Model specification using only entity fixed effects. ***, **, * denote statistical significance at the 1, 5, 10 %-level.

in this and the subsequent subsample. Yet, I still reject the strong form of efficiency. After the financial crisis of 2008 I find only little evidence against the strong form of efficiency (see the predictor industrial orders), however, I strongly reject the weak form of forecast efficiency for this subsample in contrast to previous subsamples. Due to a large difference between the adjusted R^2-statistics of the standard specifications and the adjusted R^2-statistics of the models without time fixed effects, the results of the subsample analysis also suggest a strong influence of time fixed effects on the export growth forecast error. It is striking that in the subsample "*post financial crisis*" the R^2-statistics of the model without time fixed effects are comparatively high. This indicates that the element of surprise of a given year is less important in explaining the forecast error than in previous samples, possibly because the forecast errors are the least scattered in this subsample (see Figure 2).

Regarding the subsamples of the four-quarter-ahead import growth forecasts, reported in Table 5, a similar picture emerges for the adjusted R^2-statistics. However, in contrast to their export growth counterparts, I neither find a change in weak nor in strong efficiency of these forecasts over time. Regarding unbiasedness, I find less evidence against the unbiasedness of forecasts of the "*reunified Germany*" subsample. After the financial crisis of 2008 forecasters again form biased forecasts. An explanation might be that the financial crisis led to more conservative forecasts and a tendency of underestimation among the economic research institutes (see Döpke et al., 2019b, for evidence of more cautious behavior of German professional forecasters after the financial crisis). In all subsamples German four-quarter-ahead import growth forecasts are weakly efficient, and I find evidence against the strong form of efficiency. The subsamples differ with respect to the predictors which have statistically significant explanatory power of the forecast error. In most cases the coefficients of the predictors industrial orders and leading indicator are significant. The only predictor which is insignificant in all subsamples is the trade policy uncertainty index.

Overall forecasting behavior does not change substantially over the subsamples. Yet, after the financial crisis 2008/2009, I find stronger evidence, compared to antecedent subsamples, against weak efficiency of export growth forecasts and against unbiasedness of import growth forecasts,

Table 5: Efficiency Tests of Four-Quarters-Ahead Import Growth Forecasts for Germany - Subsamples

	Subsample: Oil Crises (1971 - 1981)									
Intercept	-1.651***	0.993	-2.316***	-2.612	-2.371***	-2.457***	-2.298***	-1.450	-2.478***	-2.243***
	(0.570)	(12.251)	(0.601)	(1.757)	(0.791)	(0.676)	(0.547)	(1.425)	(0.709)	(0.600)
Industr. Orders	2.495*									
	(1.361)									
Unemp. Rate		1.709								
		(6.467)								
Busin. Climate			0.628							
			(0.669)							
Production GER				0.285						
				(1.320)						
Production G7					0.233					
					(1.013)					
Oil Price						-1.139				
						(1.526)				
Leading Indic.							2.181***			
							(0.382)			
REER								-0.328		
								(0.605)		
TPU									-0.481	
									(0.672)	
$\mathbf{e_{t-1}}$	-0.128	-0.108	-0.140	-0.119	-0.120	-0.106	-0.127	-0.104	-0.117	-0.115
	(0.099)	(0.101)	(0.106)	(0.097)	(0.099)	(0.097)	(0.088)	(0.092)	(0.100)	(0.098)
Observations	55	55	55	55	55	55	55	55	55	55
Adjusted R^2	0.900	0.893	0.895	0.893	0.893	0.894	0.917	0.894	0.894	0.896
Adjusted R^2 **(w/o time FE)**	0.048	0.070	0.302	0.097	0.095	0.358	0.037	0.008	0.056	0.028
	Subsample: West Germany (1982 - 1992)									
Intercept	3.260***	1.383	1.756***	4.917***	2.737***	1.681***	1.260	1.370	1.745***	1.750***
	(0.549)	(1.200)	(0.397)	(0.763)	(0.708)	(0.353)	(1.692)	(2.201)	(0.333)	(0.331)
Industr. Orders	2.541***									
	(0.767)									
Unemp. Rate		-0.896								
		(1.822)								
Busin. Climate			0.007							
			(0.250)							
Production GER				3.343***						
				(0.758)						
Production G7					1.400					
					(0.898)					
Oil Price						-0.200				
						(0.308)				
Leading Indic.							1.260			
							(1.419)			
REER								-0.121		
								(0.737)		
TPU									-0.059	
									(0.550)	
$\mathbf{e_{t-1}}$	-0.047	-0.054	-0.054	-0.054	-0.031	-0.058	-0.049	-0.054	-0.053	-0.054
	(0.099)	(0.108)	(0.108)	(0.097)	(0.109)	(0.106)	(0.107)	(0.108)	(0.109)	(0.108)
Observations	79	79	79	79	79	79	79	79	79	79
Adjusted R^2	0.924	0.912	0.912	0.928	0.916	0.913	0.912	0.912	0.912	0.914
Adjusted R^2 **(w/o time FE)**	0.017	0.067	0.371	-0.083	-0.004	0.251	0.201	0.102	-0.085	-0.076

Continued on next page.

Continued from previous page.

	Subsample: Reunified Germany (1994 - 2008)									
Intercept	-0.288 (1.715)	-1.564 (2.573)	-3.505*** (1.069)	-0.676 (2.768)	-1.920 (1.319)	-3.401*** (1.099)	-1.179 (1.431)	-1.726 (1.527)	-3.359*** (1.244)	-3.496*** (1.063)
Industr. Orders	2.075** (0.864)									
Unemp. Rate		-3.325 (4.236)								
Busin. Climate			-0.101 (0.431)							
Production GER				1.896* (1.134)						
Production G7					2.281** (1.133)					
Oil Price						0.194 (0.484)				
Leading Indic.							1.842** (0.799)			
REER								-1.355* (0.800)		
TPU									-0.109 (0.554)	
$\mathbf{e_{t-1}}$	-0.060 (0.075)	-0.101 (0.081)	-0.095 (0.082)	-0.074 (0.078)	-0.065 (0.078)	-0.086 (0.081)	-0.060 (0.076)	-0.065 (0.076)	-0.093 (0.081)	-0.092 (0.080)
Observations	109	109	109	109	109	109	109	109	109	109
Adjusted R^2	0.885	0.879	0.878	0.882	0.883	0.879	0.885	0.882	0.878	0.880
Adjusted R^2 **(w/o time FE)**	0.198	0.497	0.007	0.111	0.000	0.087	0.069	0.033	0.026	0.007
	Subsample: Post Financial Crisis (2010 - 2018)									
Intercept	-4.058 (3.210)	-11.862** (3.621)	-5.879*** (0.772)	-6.640*** (2.095)	-6.648** (3.067)	-6.204*** (0.869)	-4.566*** (0.711)	-5.649*** (0.998)	-5.827*** (0.842)	-5.778*** (0.832)
Industr. Orders	0.391 (0.803)									
Unemp. Rate		14.981* (8.599)								
Busin. Climate			0.978** (0.413)							
Production GER				-0.196 (0.533)						
Production G7					-0.188 (0.732)					
Oil Price						0.499** (0.232)				
Leading Indic.							1.386*** (0.356)			
REER								0.465 (1.050)		
TPU									-0.129 (0.286)	
$\mathbf{e_{t-1}}$	-0.137 (0.117)	-0.141 (0.108)	-0.127 (0.111)	-0.144 (0.115)	-0.188 (0.732)	-0.132 (0.108)	-0.118 (0.108)	-0.147 (0.119)	-0.139 (0.111)	-0.141 (0.112)
Observations	69	69	69	69	69	69	69	69	69	69
Adjusted R^2	0.847	0.850	0.855	0.847	0.847	0.851	0.855	0.847	0.847	0.850
Adjusted R^2 **(w/o time FE)**	0.400	0.360	0.161	0.428	0.583	0.215	0.136	0.140	0.128	0.142

Notes: Results are computed by means of an entity and time fixed effects regression with panel-corrected standard errors (PCSE, Beck and Katz, 1995). Dependent variable: Import growth forecast error (four-quarters-ahead). PCSE in parentheses. Entity and time fixed effects are excluded to save journal space. W/o time FE: Model specification using only entity fixed effects. ***, **, * denote statistical significance at the 1, 5, 10 %-level.

indicating a change in forecasting behavior.

4 Concluding Remarks

I have built on the literature evaluating German trade forecasts (Behrens, 2020, 2019) by means of panel regressions (Döpke and Fritsche, 2006, who evaluate GDP and inflation forecasts). Furthermore, I have contributed to the literature on the effects of economic recessions on professional forecasters (Frenkel et al., 2011; Döpke et al., 2018, 2019a). I have followed Döpke et al. (2019a), who analyze German GDP and inflation forecasts, and have estimated a fixed effects panel regression using Beck and Katz (1995) panel-corrected standard errors. To this end, I have used a novel data set on German trade forecasts for the years 1971 to 2019 of eight leading German professional forecasters. To analyze possible changes in forecasting behavior, I have estimated the fixed effects model for subsamples after incisive economic events in recent German history, namely the oil price shocks in the early and late 1970s, German reunification, and the financial crisis of 2008/2009.

I find that all analyzed German trade forecasts in the full sample, i.e., export and import growth as well as two- and four-quarter-ahead, are in line with the concept of weak efficiency. In other words, the lagged forecast error does not have explanatory power for the current forecast error. Furthermore, the shorter-term forecasts are not biased, whereas I find evidence against the unbiasedness for both types of longer-term trade forecasts. I reject the strong form of efficiency for all forecasts in the full sample. Predictors with explanatory power in most specifications are mainly the OECD leading indicator and industrial orders for Germany. Time fixed effects, which can be interpreted as the element of surprise of the year for which a forecast was formed, play a crucial role in explaining the forecast errors.

Overall, the results for the subsamples with respect to strong efficiency resemble the ones for the full sample, which is in line with recent research on possible changes in the behavior of professional forecasters after the financial crisis (Döpke et al., 2019a). However, there are differences regarding the weak form of efficiency of the export growth forecasts. Especially after the financial crisis, I strongly reject weak efficiency of these forecasts. Regarding import growth forecasts, a bias is less of an issue after German reunification and before the financial crisis. Before and after this time period, I find strong evidence against the unbiasedness of the longer-term import growth forecasts. In summary, I find changes in forecasting behavior of trade forecasters after the financial crisis, namely regarding weak efficiency and unbiasedness. A possible explanation is a tendency to more conservative forecasts after the financial crisis 2008/2009 as has been reported by Döpke et al. (2019b).

In future research, it will be interesting to further analyze this change of forecasting behavior and the reported bias of the longer-term trade forecasts. The latter effect could be explained by further analyses of a possible asymmetry of the forecasters' loss functions (building on research by Behrens, 2019) or of possible behavioral biases in trade forecasts (see, e.g., Ito, 1990, on wishful expectations). It will also be interesting to see, if the change in forecasting behavior after the financial crisis can be confirmed for other macroeconomic aggregates as well as for the period after the Covid-19 pandemic and the associated economic crisis.

References

Arellano, M. / *Bond,* S. (1991): "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations", The Review of Economic Studies 58(2), 277–297.

Bailey, D. / *Katz,* J. N. (2011): "Implementing panel-corrected standard errors in R: The pcse package", Journal of Statistical Software, Code Snippets 42(1), 1–11.

Beck, N. / *Katz,* J. N. (1995): "What to do (and not to do) with Time-Series Cross-Section Data", The American Political Science Review 89(3), 634–747.

Behrens, C. (2019): "A nonparametric evaluation of the optimality of German export and import growth forecasts under flexible loss", Economies 7(3), 93.

– (2020): "Evaluating the joint efficiency of German trade forecasts - A nonparametric multivariate approach", Applied Economics 52(34), 3732–3747.

Behrens, C. / *Pierdzioch,* C. / *Risse,* M. (2018a): "A test of the joint efficiency of macroeconomic forecasts using multivariate random forests", Journal of Forecasting 37(5), 560–572.

– (2018b): "Testing the optimality of inflation forecasts under flexible loss with random forests", Economic Modelling 72, 270–277.

– (2020): "Do German economic research institutes publish efficient growth and inflation forecasts? A Bayesian analysis", Journal of Applied Statistics 47(4), 698–723.

Behrens, K. / *Corcos,* G. / *Mion,* G. (2013): "Trade crisis? What trade crisis?", Review of Economics and Statistics 95(2), 702–709.

Bems, R. / *Johnson,* R. / *Yi,* K.-M. (2010): "The role of vertical linkages in the propagation of the global downturn of 2008", IMF Economic Review 58(2), 295–326.

Berman, N. / *Martin,* P. (2012): "The vulnerability of sub-saharan Africa to the financial crisis: The case of trade", IMF Economic Review 60(3), 329–364.

Caldara, D. / *Iacoviello,* M. / *Molligo,* P. / *Prestipino,* A. / *Raffo,* A. (2019): "The economic effects of trade policy uncertainty", manuscript presented at the 91st meeting of the Carnegie-Rochester-NYU Conference on Public Policy.

Campos, N. F. / *Fidrmuc,* J. / *Korhonen,* I. (2019): "Business cycle synchronisation and currency unions: A review of the econometric evidence using meta-analysis", International Review of Financial Analysis 61, 274–283.

Capistrán, C. (2008): "Bias in Federal Reserve inflation forecasts: Is the Federal Reserve irrational or just cautious?", Journal of Monetary Economics 55(8), 1415–1427.

Chor, D. / *Manova,* K. (2012): "Off the cliff and back? Credit conditions and international trade during the global financial crisis", Journal of International Economics 87(1), 117–133.

D'Agostino, A. / *Modugno,* M. / *Osbat,* C. (2017): "A global trade model for the Euro Area", International Journal of Central Banking 13(4), 1–34.

Döhrn, R. / *Schmidt,* C. M. (2011): "Information or Institution? On the determinants of forecast accuracy", Journal of Economics and Statistics 231(1), 9–27.

Döpke, J. / *Fritsche,* U. (2006): "Growth and inflation forecasts for Germany - A panel-based assessment of accuracy and efficiency", Empirical Economics 31(3), 777–798.

Döpke, J. / *Fritsche,* U. / *Müller,* K. (2019a): "Has macroeconomic forecasting chnaged after the Great Recession? Panel-based evidence on forecast accuracy and forecaster behaviour from Germany", Journal of Macroeconomics 62, 103–135.

Döpke, J. / *Fritsche,* U. / *Pierdzioch,* C. (2017): "Predicting recession with boosted regression trees", International Journal of Forecasting 33(4), 745–759.

Döpke, J. / *Fritsche,* U. / *Waldhof,* G. (2019b): "Theories, techniques and the formation of German business cycle forecasts", Journal of Economics and Statistics 239(2), 203–241.

Döpke, J. / *Müller,* K. / *Tegtmeier,* L. (2018): "The economic value of business cycle forecasts for potential investors - Evidence from Germany", Research in International Business and Finance 46, 445–461.

Drechsel, K. / *Scheufele,* R. (2012): "The performance of short-term forecasts of the German economy before and during the 2008/2009 recession", International Journal of Forecasting 28(2), 428–445.

Durusoy, S. / *Sica,* E. / *Beyhan,* Z. (2015): "Economic crisis and protectionism policies: The case of the EU countries", International Journal of Humanities and Social Science 5(6), 57–68.

Eaton, J. / *Kortum,* S. / *Neiman,* B. / *Romalis,* J. (2016): "Trade and the global recession", American Economic Review 106(11), 3401–3438.

Engelke, C. / *Heinisch,* K. / *Schult,* C. (2019): "Effects of external assumptions on forecast errors", IWH Discussion Paper, forthcoming.

Evenett, S. J. (2009): "What can be learnt from crisis-era protectionism? An initial assessment", Business and Politics 11(3), 1–26.

Frale, C. / *Marcellino,* M. / *Mazzi,* G. L. / *Proietti,* T. (2010): "Survey data as coincident or leading indicators", Journal of Forecasting 29(1-2), 109–131.

Frenkel, M. / *Lis,* E. M. / *Rülke,* J.-C. (2011): "Has the economic crisis of 2007-2009 changed the expectation formation process in the euro area?", Economic Modelling 28(4), 1808–1814.

Freund, C. (2009): The great trade collapse: Causes, consequences, and prospects, CEPR, chap. The trade response to global downturns. 59–70.

Gaibulloev, K. / *Sandler,* T. / *Sul,* D. (2014): "Dynamic panel analysis under cross-sectional dependence", Political Analysis 22(2), 258–273.

Grimme, C. / *Lehmann,* R. / *Noeller,* M. (2018): "Forecasting imports with information from abroad", CESifo Group No. 7079.

Guichard, S. / *Rusticelli,* E. (2011): "A dynamic factor model for world trade growth", OECD Economics Department Working Papers No. 874.

Heilemann, U. / *Stekler,* H. O. (2013): "Has the accuracy of German macroeconomic forecasts improved?", German Economic Review 14(2), 235–253.

Heinisch, K. / *Scheufele,* R. (2018): "Bottom-up or direct? Forecasting German GDP in a data-rich environment", Empirical Economics 54, 705–745.

Holden, K. / *Peel,* D. A. (1990): "On testing for unbiasedness and efficiency of forecasts", The Manchester School 58(2), 120–127.

Huber, P. J. (1967): "The behavior of maximum likelihood estimates under nonstandard conditions", Proceedings of the fifth Berkeley symposium on mathematical statistics and probability 1(1), 221–233.

Ito, T. (1990): "Foreign exchange rate expectations: Micro survey data", The American Economic Review 80(3), 434–449.

Keane, M. P. / *Runkle,* D. E. (1990): "Testing the rationality of price forecasts: New evidence from panel data", The American Economic Review 80(4), 714–735.

Kee, H. L. / *Neagu,* C. / *Nicita,* A. (2013): "Is protectionism on the rise? Assessing national trade policies during the crisis of 2008", Review of Economics and Statistics 95(1), 342–346.

Kirchgässner, G. / *Müller,* U. K. (2006): "Are forecasters reluctant to revise their predictions? Some German evidence", Journal of Forecasting 25(6), 401–423.

Kmenta, J. (1986): "Elements of econometrics", Macmillan, New York, second edition.

Lehmann, R. (2015): "Survey-based indicators vs. hard data: What improves export forecasts in Europe?", ifo Working Paper No. 196.

Levchenko, A. A. / *Lewis,* L. T. / *Tesar,* L. L. (2010): "The collapse of international trade during the 2008-09 crisis: In search of the smoking gun", IMF Economic Review 58(2), 214–253.

MacKinnon, J. G. / *White,* H. (1985): “Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties”, Journal of Econometrics 29(3), 305–325.

Mincer, J. A. / *Zarnowitz,* V. (1969): “The evaluation of economic forecasts”, in: Economic forecasts and expectations: Analysis of forecasting behavior and performance, ed. by Mincer, J. A., Economic forecasts and expectations: Analysis of forecasting behavior and performance, 3–46, National Bureau of Economic Research, New York.

Nickell, S. (1981): “Biases in dynamic models with fixed effects”, Econometrica 49(6), 1417–1426.

Öller, L.-E. / *Barot,* B. (2000): “The accuracy of European growth and inflation forecasts”, International Journal of Forecasting 16(3), 293–315.

Pain, N. / *Lewis,* C. / *Dang,* T.-T. / *Jin,* Y. / *Richardson,* P. (2014): “OECD forecasts during and after the financial crisis: A post mortem”, OECD Economics Department Working Papers, No. 1107.

Parks, R. W. (1967): “Efficient estimation of a system of regression equations when disturbances are both serially and contemporaneously correlated”, Journal of the American Statistical Association 62(318), 500–509.

R Core Team (2020): R: A language and environment for statistical computing, R: version 3.6.1. R Foundation for Statistical Computing, Vienna, Austria, `https://www.R-project.org/`.

Sinclair, T. / *Stekler,* H. O. / *Müller-Dröge,* H. C. (2016): “Evaluating forecasts of a vector of variables: A German forecasting competition”, Journal of Forecasting 35(6), 493–503.

Stock, J. H. / *Watson,* M. W. (2007): “Introduction to Econometrics”, Pearson Education Inc., second edition.

Timmermann, A. (2007): “An evaluation of the *World Economic Outlook* forecasts”, IMF Staff Papers 54(1), 1–33.

Wang, Y. / *Lee,* T.-H. (2014): “Asymmetric loss in the greenbook and the survey of professional forecasters”, International Journal of Forecasting 30(2), 235–245.

White, H. (1980): “A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity”, Econometrica 48(4), 817–838.