N° 121 ÉCONOMIE ET STATISTIQUES
Mai 2021 WORKING PAPERS DU STATEC

Using panel regressions to evaluate and update economic forecasts

Author :
Michel Geller (STATEC)

Abstract

In this applied working paper, we use a panel of IMF's historical forecasts for annual real GDP in order to update those forecasts in real time, shortly after their publication. We regress forecast errors on past forecast errors and selected leading indicators in a pseudo real-time forecasting exercise. We augment the IMF's forecasts with predicted forecast errors and show that our procedure reduces it by up to 14%, one month after the IMF forecast is released. We nd that past forecast errors improve prediction accuracy by more than leading indicators do. In addition, this procedure works especially well for open economies and in simpler models like pooled OLS.

Using panel regressions to evaluate and update economic forecasts

Michel Geller

April 2, 2021

Abstract

In this applied working paper, we use a panel of IMF's historical forecasts for annual real GDP in order to update those forecasts in real time, shortly after their publication. We regress forecast errors on past forecast errors and selected leading indicators in a pseudo real-time forecasting exercise. We augment the IMF's forecasts with predicted forecast errors and show that our procedure reduces it by up to 14%, one month after the IMF forecast is released. We find that past forecast errors improve prediction accuracy by more than leading indicators do. In addition, this procedure works especially well for open economies and in simpler models like pooled OLS.

1 Introduction

In order to make macroeconomic policy decisions, today's economies rely on accurate and timely GDP forecasts. Planning governmental budget decisions, anticipating times of crisis and organising appropriate countermeasures, are all based on reliable forecasts on the future state of the economy. Hence, one of the most important macroeconomic tasks is the fall year-ahead forecast, as it is the basis of next year's budget decisions.

In economic forecasting, 2 groups of tasks are of particular importance: the a priori side leading to the creation of the forecast, choosing the right econometric models, assumptions and variables. And then there is the posterior task of controlling the accuracy of the forecasts and realisations and adapting the model, after the observations have been published. Naturally, each of these tasks can be broken down into smaller questions such as choosing the right predictor variables or which observations are suited to benchmark the model.

The present study acts at the connection of these two parts. A panel regression is defined, explaining real GDP forecast errors by early indicator variables in order to capture and understand lacking information in the original forecasts. In a second step, these equations are used to anticipate these forecast errors for future rounds. By adding the anticipated forecast error to the original forecast, a new, adjusted forecast is created. As the early indicator variables are typically of monthly frequency, this may be reproduced at different horizons, such as at the moment the IMF releases their original forecasts, and during the months shortly after their forecast is published. In this work, the method targets only fall year-ahead forecasts, but could be extended to other episodes such as the spring year ahead or the spring in-year forecast. Typically forecasts on long periods, such as those 15 months ahead by the IMF, profit less from monthly frequency equations as inter-annual fluctuations are less important on long forecasting horizons. The main aim is to improve anticipation of major downturns such as the Great Recession. To achieve this, pseudo-historical series are assembled to see how a better selection of indicators would have improved international year-ahead GDP forecasts a priori. A similar study for quarterly data has been done by Chen and Rancière.

Traditional forecasting models, which are used to make predictions up to the medium term, rely on economic theory in order to produce predictions. Such models include traditional large scale macroeconometric models such as the OECD's NiGEM (Turner 2016), which are widely used among central banks and national and international forecasting divisions such as DG-ECfin at the European Commission. Often, these theory-based models lack channels of financial or survey data (Del Negro et al. 2013) which are normally more important in short term models. On the other hand, short-term forecasting models such as dynamic factor models are not restrained to theory relevant data sources as they identify signals from vast amounts of data, however their precision is rather low for forecasts over the horizon of one year and more (see Banerjee and Marcelino).

Despite its forecasting aim, the underlying work relies heavily on the posterior side of the forecast, namely forecast errors. Year-ahead GDP forecasting is a demanding task, accompanied by various pitfalls and sources of errors. There are many recent forecast evaluations, such as Pain and Lewis 2015 on OECD forecasts, Fioramanti et al. 2016 on European Commission Forecasts or Genberg and Martinez 2014 on IMF forecasts. They find that year-ahead forecasts typically fail to anticipate crisis appropriately (especially the Great Recession), leading to large forecast errors at the beginning of recessions and less so at the recovery.

The IMF provides a panel of historic yearly forecasts and historic observations from the World Economic Outlook (WEO). This panel serves as the basis of this work. It contains detailed time series vintages for over 180 countries for the last 25 years. Out of it, 16 OECD member states have been selected, providing similar or identical predictor time series. As the time dimension is fairly short, a panel-model gives more observations and estimation stability. This panel structure allows for out-of-sample forecast adjustments. Some of the equations show significant forecast error reduction, a pooling equation for instance is able to reduce forecast errors at a 14 months horizon on average by 0.18 percentage points, which reduces the original IMF forecast error by 14 percent.

It is not the intend of this paper to find optimal predictive variables, it relies on standard early indicator variables such as those examined by the OECD in Astolfi et al. 2016 and used in Chen and Rancière 2017. The variables cover the housing market, the stock market, public sector risk premia and consumer surveys.

The results show an average reduction of forecast errors. Globalised open economies such as Denmark, Luxembourg or the United Kingdom tend to profit most from the equations, while others such as Belgium and Australia do not profit. The panel leaves a lot of room for adaptation to specific needs, and a lot of future research needs to be done.

The paper is organised as follows. Section 2 describes the underlying data, section 3 describes the model, section 4 presents the results, only to conclude in section 5.

2 Data

2.1 Year ahead forecast and observed GDP

All of the presented GDP data, i.e. forecasts and observations come from the World Economic Outlook database. This database contains some 180 countries and gives forecasts and observation vintages from 1988 up to today. The IMF database is updated twice a year in spring and fall. The forecasts reach up to an horizon of 5 years, while observations are available for $t-1$ and $t-2$, where t refers to the year of the corresponding publication. E.g. GDP_{t-1} observed in autumn t refers to real GDP for the year 2019, observed in Autumn 2020, GDP_{t+1} refers to real GDP in 2021 and is unobserved in the same period. To prevent methodological issues, all countries are members of the organisation for economic cooperation and development. This allows to use a selection of economic variables (OECD Main Economic Indicators). Further, OECD members have similar states of economic development, implying a certain homogeneity across variables. In a later chapter, the panel forecast is applied to a sub-sample consisting of European countries, allowing for a very homogeneous sample. Hence, the panel selection for this work is confined to 16 countries.

The yearly GDP estimation for the past year, GDP_{t-1} observed in autumn t, is done by the national institutes before it is sent to the IMF. However, in some cases the yearly estimates for $t-1$ are not available when the IMF starts its forecasting procedure. In these cases, usually the last available estimation, based on quarterly accounts is used by the IMF. In the current work we do not check for this issue and take the data as given and of homogeneous origin. In order to have the smallest time interval between forecasts and observations, the first observation, which is released around September $t+1$, is used as yearly observation. This approach is the standard approach for the IMF, OECD and European Commission to evaluate forecast errors. Lastly, it is important to acknowledge, that in autumn of year t only a GDP estimation up to year $t-1$ is available, so an auto-regressive approach towards year-ahead GDP forecasting would start at a lag of 2 years or required iterative forecasting. Forecast errors are defined as: $Forecast\_error_t = Observation_{t,t+1} - Forecast_{t,t-1}$. Where the forecast error from year t is composed of the observation realised in year $t+1$ and the year ahead forecast realised in year $t-1$.

The descriptive statistics on GDP, the forecasts and forecast errors, are presented for an international sample, consisting of Australia, Belgium, Canada, Denmark, the Euro Area, Finland, France, Germany, South Korea, Luxembourg, the Netherlands, New Zealand, Spain, Sweden, the United Kingdom and the United States. While the lowest GDP growth rates observed in the sample are -8% (realised by Finland in 2009), the lowest forecast was only -1.32%, targeting Spain in 2013. All in all the forecast errors are significantly biased to the downside, i.e. forecasts are overly optimistic. Mean absolute forecast errors are about 1.27 percentage points.

2.2 Early Indicators

As mentioned in the introduction, this paper is not a work on finding optimal leading indicators for recessions and times of large GDP volatility. Instead the estimations rely on leading indicators examined by prominent international institutions like the OECD in Astolfi et al. 2016, and the Federal Reserve Bank of New York in Estrella and Mishkin 1996. Some of the most common early indicators have been selected and implemented with considerable success. However to avoid multicolinearity and keep the degrees of freedom up, 4 early indicator variables plus a constant found

Figure 1: Real GDP: Forecasts and Observations

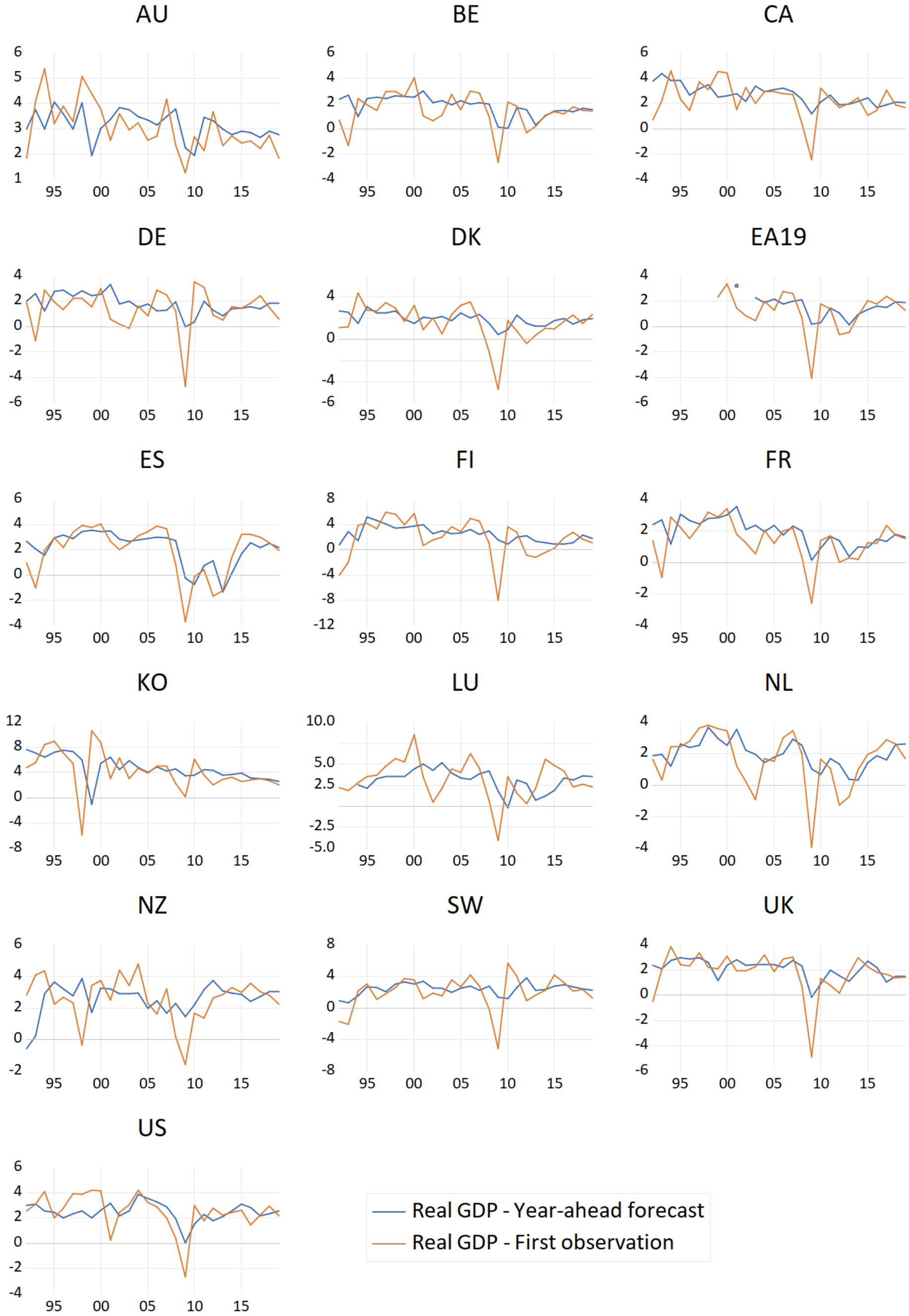

Figure 1: Comparing year ahead GDP forecasts and first observations. Large differences appear during economic downturns such as the Great Recession (2007-2009) or the early 2000s recession (2000-2001).

Table 1: Descriptive statistics

	Mean	Maximum	Minimum	Std. Dev.	Mean Absolute Forecast Error
Observed Real GDP	2.05	10.66	-8.02	2.15	
Real GDP year-ahead forecast	2.40	7.61	-1.32	1.21	
Forecast Error	-0.37***	11.66	-11.84	-11.84	1.33

*** Significant at 0.01-level
Sample: 1992-2019
Data: IMF World Economic Outlook
Forecast Error = Observation - Forecast

Table 1: Real GDP: Observations, forecasts and forecast errors: Standard deviations and extreme values show higher volatility for GDP observations than for the respective forecasts, resulting in forecast errors. Further, the forecast errors are significantly biased to the upside.

their way into the regressions. In line with the limitations of the WEO sample, the horizon of all regressors is restricted to 1991 to 2019.

The early indicators used are residential housing permit transactions, national stock market indices, risk premia and consumer surveys. Besides residential housing permits which is only available in quarterly frequency for some countries, all variables are in monthly frequency. Government bond risk-premia represent the difference between a country's 10 year government bond yield and a country's 2 year government bond yield. The most prominent way to include government risk-premia would be to take the term spread, the difference between a country's 3 month and 10 year Treasury. However due short data on the 3 month government bonds, the above method was preferred in order to keep the data-set from shrinking. Volatility indices such as VSTOXX or VIX have not been included as they are not available for each country in the sample and besides the VIX, these indicators have not been published for a long period. Variables were rendered stationary by first differencing and standardized. Whenever possible the variables were taken from the OECD Main Economic Indicators (MEI) database. For small countries such as Luxembourg, not all variables are representative for the economy or simply are to volatile to give good results. In this case the variables are replaced by the respective Euro Area variables. A notable example, is the Luxembourgish stock market, which is composed of 9 companies in total, of which 30% of the shares belong to steel companies. As steel is only a minor business in the Luxembourgish economy this is not representative of its structure. Hence the Eurostoxx50 is used as a proxy for Luxembourgish stock market information. The appendix gives a complete listing of the series (see table 5).

To align monthly variables with yearly real GDP, they were time-aggregated by taking 3 month averages of data available at the time of the forecast. This way the most recent past of each variable is taken into consideration, while outliers are reduced by taking the 3 month average.

3 Method

3.1 Panel models

The basic idea of this work is to evaluate and track annual forecast errors. First, the forecast errors are regressed on the early indicator variables, explaining the historic forecast errors by deviations in these non structural variables. Then the panel regression allows to forecast the forecast errors, in order to anticipate future deviations. The process resembles the Mincer-Zarnowitz panel regression (Mincer-Zarnowitz 1969), where biases in forecasts are analysed. However, in the underlying work the focus is less on potential biases and more on omitted information in the forecast. The main contribution of the paper is the ability to track the IMF forecast, based on the evolution of the early indicator variables, in the months after the publication of the initial forecast. This is done by a set of panel equations.

Three equations are the main engines of this work: a pooled ordinary least squares equation $FCERROR_t = c+\beta*X_\tau+u_t$, a standard random effects equation $FCERROR_{i,t} = c+\beta*X_{i,\tau}+u_{i,t}$ and a fixed effects equation $FCERROR_{i,t} = c + \beta * X_{i,\tau} + u_{i,t} + fe_i$. In a fixed effects model, country specific constants are assumed to be non-random parameters. This somewhat technical assumption can reduce the model efficiency which could result in less precise forecasts. A random effects model assumes that country specific constants are random parameters. If this assumption holds the forecasts will be more efficient than in a fixed effects model. However, if it is violated the model forecasts will be biased. Nonetheless, even if biased the gain in estimation efficiency could make them more precise (on average). An extreme example of such an occurrence are pooled OLS models that feature one constant term equal for all countries. The reduction in the number of parameters estimated might outweigh the single-constant bias making this model perform best in terms of forecasting. Keeping these theoretical considerations in mind we estimate all three and compare their forecasting performance. $FCERROR_{i,t}$ is the GDP forecast error as defined earlier, i indicating the country while t being the respective year. $X_{i,\tau}$ is the set of regressors, while the vector β contains the regression coefficients. The regressors are aggregated to yearly frequency up from monthly frequency. The time parameter τ is different to t as it follows a monthly frequency and it may be adapted to various time horizons. If this regression is run at the moment of the original forecast, τ is usually equal to $t-15$ months, i.e. the regressors are aggregated into yearly data points ending 15 months before the end of t. However τ maybe adapted to later points in time, i.e. $\tau = t - 10$ to run a forecast verification 10 months before the end of the year, at which the original forecast was aimed. In this case the regressors are aggregated over yearly cycles with the last month ending 6 months before the end of the year in question. c is a constant accounting for a possible bias. $u_{i,t}$ is the composite error, for country i and year t. Finally fe_i is the fixed effect for country i.

In order to evaluate the performance of auto regressive regressors, a pseudo lagged forecast error (abbreviation PLFCE) has been added in a second set of equations. The pseudo lagged forecast error is defined as the difference between the year-ahead and in year forecast of the previous year. As GDP is observed with a lag of 1 year, a true lagged forecast error would not be relevant for out of sample forecasting nor ordinary forecasting. Dynamic panel equations, come at the price of a bias (Nickell 1981), especially for small t. In this paper we do not go further into this and expect t to be large enough, not to have a strong bias. A possible solution to this issue is Arellano-Bond estimation.

In these simple panel and pooled regressions, the equations show significant cross section dependence when applying the Pesaran cross section dependence test (Pesaran 2004). Considering the fact that all of our regressors follow the business cycle this should not come as a surprise. A similar reasoning applies to the heteroscedasticity in the regressors, both crisis in the sample figure as major disturbances to our regressors as well as the explanatory variables. Applying White cross section errors (White 1980) should improve the theoretical significance and consistency of the panel. However it does not remove cross section dependence. Further, GLS based cross-section weights (see for example Pesaran 2015) allow for different residual variances per cross section. A promising extension is a Seemingly Unrelated Regressions model (Zellner 1962). However, it is constrained by needing longer balanced time series, so it is only applicable in a sub-sample (see table 4). Unfortunately it is not solving the problem of cross section dependence. Nevertheless, this is a forecasting exercise, so the out-of-sample forecasting properties and their consistency analysed in the following chapter matter most.

4 Results

The results part is separated into three sub-chapters: regression results, out of sample forecast results and an analysis of possible extensions to the model.

Figure 2: Rolling regressions of a pooled model

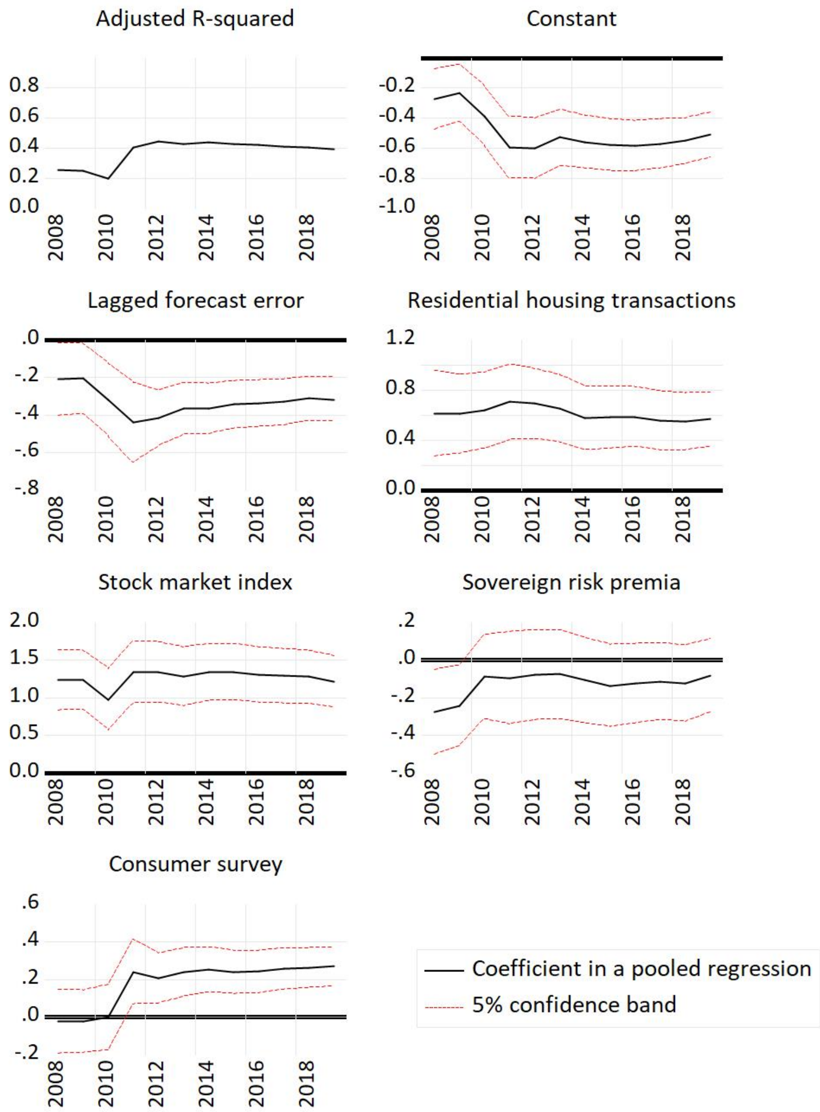

Figure 2: Rolling regression estimates of a pooled model with a pseudo lagged forecast error as an additional regressor. Rolling regressions are anchored in 1992.
Reading example: The leftmost, or 2008 point in each graph is the corresponding coefficient of a pooled regression using data from 1992 to 2008 and including all of the depicted variables. If both confidence bands are either positive or negative, for an estimated coefficient, then this coefficient is significant at the 5% level. For example sovereign risk premia are a significant regressor with negative sign in a pooled regression using data from 1992 to 2008. Albeit including 2010 into the data, the sovereign risk premia is not significant anymore.

4.1 Regression

In this section, we analyse stability of the regression models by plotting regression coefficients against time, using extending window samples. Coefficients are quite stable after data includes the Great Recession, implying a stable and relationship between the regressors and real GDP forecast errors.

The panel regressions contain 28 years, from 1992 to 2019, and a total of 368 unbalanced observations. In order to check for the stability of the equations and the corresponding coefficients, rolling regressions were performed for each equation. Due to short time series, we are using an extending window rolling regression, anchored in 1992. Hence the regressions prior the crisis have considerably less observations than at the end of the sample.

The results are shown in graphs where the coefficient is displayed per year at the end of the respective period. 5% confidence intervals are added to show significance of the estimates, as long as both confidence intervals are either negative or positive, the coefficient is significant at the 5% level. When looking at the results for the pooled, fixed effects and random effects models with a pseudo lagged forecast error in figure 2 (separate rolling regression results may be found in the appendix), coefficients are quite constant over time. As most variables, the constant is strongly affected by the Great Recession. As the IMF predicted quite stable growth for 2008 and 2009, seen in figure 1, the observed negative growth rates for these years affected the constant to the downside. Considering the regressors, the housing permit transactions and the stock market, react to the the crisis in a similar way, especially for equations with less lead. For year ahead leads (16 to 12 months), consumer surveys became significant only after including the Great Recession in the rolling regressions. Government risk premia show the least significance in the regressions, being partially significant for a lead of 16 to 15 months. Considering the signs of the coefficients, most are in line with theory, i.e. positive signs for residential housing transactions, stock market indices and surveys, while negative signs for the constant, lagged forecast error and government risk premia. However, government risk premia present an other outlier here. At a lead of 16 months, they are significant regressors with a negative sign, while over closer leads, the significance fades, and the sign even changes into the positive.

Among all equations, the adjusted R-squared improves dramatically after the crisis, jumping to values between 0.4 and 0.6 at the end of the sample. The explanatory variables used in the regression are sensitive to the Great Recession, especially residential housing transactions and stock market indices, so after the Great Recession is included in the time series (i.e after 2009 in the rolling regression), the fit of the equations rises.

The results have been reproduced for a sub-sample consisting of the European countries within the sample. Those results are very similar and may be found in the appendix.

4.2 Out of sample benchmark

4.2.1 Out of sample performance at different horizons

The out of sample performance of the equations is measured by the mean absolute forecast errors and the corresponding relative forecast error, compared to the original IMF forecasts. Using both IMF-forecasts and -historical observations allow for a straightforward comparison of the panel forecasts with those by the IMF.

The benchmark setting is straightforward. For each year to forecast, a panel regression including all data available up to a certain horizon before the end of the year in question is estimated. The resulting point forecast is compared to the original IMF forecast. The difference in accuracy is presented by 2 measures: On one hand the difference of the mean absolute forecast errors, where a positive sign indicates a lower forecast error than the original IMF forecast, while a negative sign indicates a higher forecast error than the original IMF forecast (i.e $MAFE_{IMF} - MAFE_{panelmodel}$, where MAFE stands for Mean Absolute Forecast Error). Second the ratio of the respective panel mean absolute forecast errors over the original IMF mean absolute forecast errors. Here a number smaller than one indicates a smaller forecast error by the panel estimation (i.e. $MAFE_{IMF}/MAFE_{panelmodel}$).

Table 2 shows the difference of average absolute forecast errors as well as the relative forecast error with regard to the original IMF forecast.

At the 16 and 15 months lead none of the presented equations significantly reduces forecast errors. However, those equations using the pseudo lagged forecast errors are more accurate, 10 percentage points on average. 15 months before the end of the year is of particular interest, as at this moment most data used was available to the IMF as well. Residential housing permit transaction mark an exception to this as they are released with a slight lag of 2 to 3 months. At a 14 months lead, which

Table 2: Absolute forecast performance

Mean absolut forecast error difference, with regard to the IMF autumn forecast, in percentage points:

Model:	16 months lead	15 months lead	14 months lead	13 months lead	12 months lead	11 months lead	10 months lead
Pool	-0.07	-0.04	0.04	0.13*	0.2***	0.28***	0.32***
Random	-0.07	-0.04	0.04	0.13*	0.2**	0.28***	0.3***
Fixed	-0.1*	-0.08	0.01	0.09	0.15*	0.23***	0.26***
Pseudo Lagged FCError	0.05	0.05	0.05	0.05	0.05	0.05	0.05
Pool with pseudo lagged FCE	0.04	0.08	0.18***	0.21***	0.23***	0.3***	0.32***
Random with pseudo lagged FCE	0.04	0.08	0.18***	0.21***	0.23***	0.29***	0.32***
Fixed with pseudo lagged FCE	0.04	0.06	0.14*	0.15*	0.16*	0.24***	0.25***

Out of sample benchmarks using data from 1992 onwards, first out of sample realisation: 2008, last realisation: 2019
Pseudo lagged FCE defined as *inyear forecast - year ahead forecast(-1)*
*** Significance at 0.01-level // ** Significance at 0.05-level // * Significance at 0.1-level

Table 2: This table compares the results of different panel equations with the original IMF forecasts. Mean absolute forecast errors of the original IMF forecasts are subtracted by the corresponding panel equation MAFEs. Each equation is run at different leads, from 16 months before the end of the year to 10 months before the end of the year. Hence, a 15 months lead corresponds to the moment of publication of the IMF forecasts.
Reading example: A pooled equation with pseudo lagged forecast error as an additional regressor, targeting real GDP with 14 months lead, has on average a 0.18 percentage points lower forecast error than the original IMF forecast for the period 2008 to 2019.

coincides with a forecast produced in early November, the equations using the pseudo lagged forecast errors significantly reduce forecast errors. Pooled and Random effects equations reduce the forecast error by 0.18 percentage points, equivalent to 86% of the original forecast error by the IMF. Fixed effects realises a 0.14 percentage points reduction of forecast errors, equivalent to 89% of the original forecast error. At smaller leads, the forecast error is reduced by up to 0.32 percentage points or 0.73 percent of the original IMF forecast. This does not surprise as a lot more information concerning the upcoming year is available at smaller leads. Considering the equations without a lagged forecast error, the gains are significant from the 13 months lead on-wards.

Table 3: Relative forecast performance

Relative forecast errors with regard to the IMF autumn forecast, ratio:

Model:	16 months lead	15 months lead	14 months lead	13 months lead	12 months lead	11 months lead	10 months lead
Pool	1.06	1.04	0.97	0.90	0.83	0.77	0.74
Random	1.06	1.04	0.97	0.90	0.84	0.77	0.75
Fixed	1.08	1.07	1.00	0.93	0.88	0.81	0.79
Pseudo Lagged FCError	0.96	0.96	0.96	0.96	0.96	0.96	0.96
Pool with pseudo lagged FCE	0.97	0.94	0.86	0.83	0.81	0.76	0.73
Random with pseudo lagged FCE	0.97	0.94	0.86	0.83	0.81	0.76	0.74
Fixed with pseudo lagged FCE	0.97	0.95	0.89	0.88	0.87	0.81	0.79

Out of sample benchmarks using data from 1992 onwards, first out of sample realisation: 2008, last realisation: 2019
Pseudo lagged FCE defined as *inyear forecast - year ahead forecast(-1)*

Table 3: Analogous to table 2, this table compares the results of different panel equations with the original IMF forecasts. The figures represent the ratio of the corresponding panel equation MAFEs divided by the MAFE of the original IMF forecasts. Each equation is run at different leads, from 16 months before the end of the year to 10 months before the end of the year. Hence, a 15 months lead corresponds to the moment of publication of the IMF forecasts.
Reading example: A pooled equation with pseudo lagged forecast error as an additional regressor, targeting real GDP with 14 months lead, has on average 0.86 times the forecast error of the original IMF forecast for the period 2008 to 2019.

Figure 3: Mean forecast error reduction for all countries, pooled with lagged forecast errors

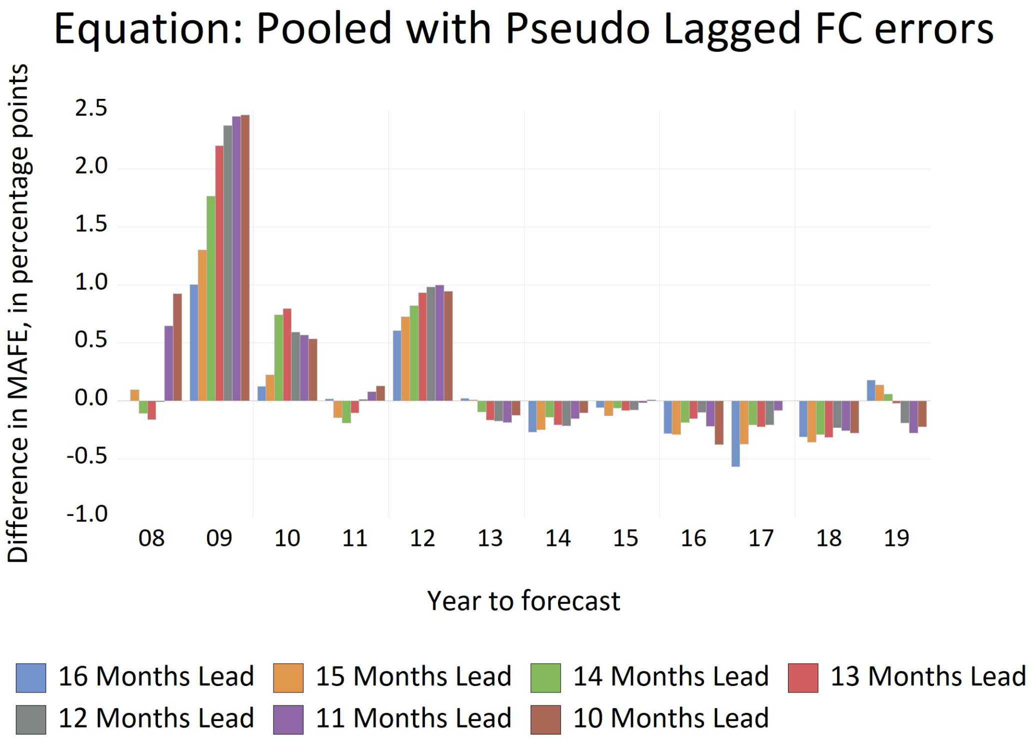

Figure 3: Mean forecast error reduction by year to forecast. Out of sample and for different forecasting horizons.
Reading example: For the 2009 year ahead forecast, produced in fall 2008, 15 months before the end of 2009, the pooled equation may reduce the original IMF forecast by 1 percentage point.

4.2.2 Year-by-year evaluation

The year to year assessment is more volatile. Figure 3 shows the average difference of absolute forecast errors by year to forecast. Through the years of the Great Recession and the European debt crisis, the panel method performs well. In 2009 the panel estimation gives 0.7 to 1 percent lower forecast errors when comparing the 16 months lead to the original IMF forecast issued with 15 months lead. At 10 months lead, the out of sample forecast error is up to 2.5 percent lower. Similarly, the set of regressors performs well in the European debt crisis, reducing the forecast error between 0.5 and 1 percentage points on average. As mentioned earlier, this might be due to the set of regressors, being composed mainly of drivers of these 2 recessionary episodes. When looking at the time after the European debt crisis, the tide turns. Between 2013 and 2019, none of the presented equations is able to persistently challenge the IMF forecast. Now, forecast errors are between 0.1 percentage points (pseudo lagged forecast error) and 0.3 percentage points (fixed effects panel regression) higher than those from the original IMF forecast.

4.2.3 Individual country performance

On the country-level large differences become apparent. Again, the dispersion is high, countries profit differently from the equations. Among those profit the least, Australia, Belgium, Canada and

Figure 4: Relative forecast errors, Luxembourg

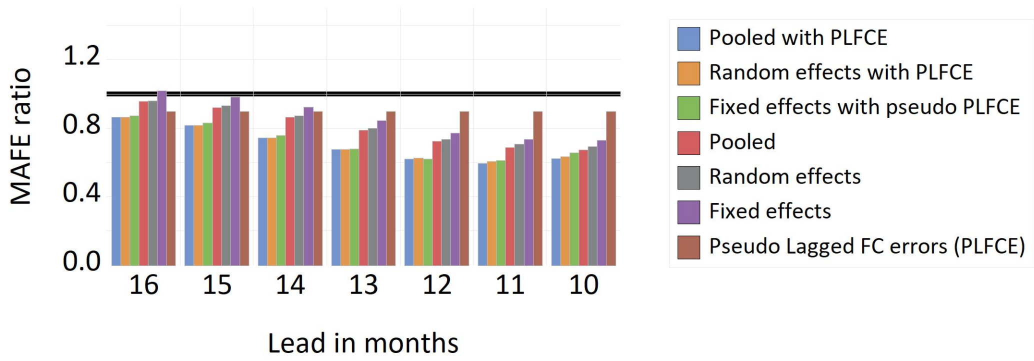

Figure 4: Relative forecast error for Luxembourg by forecast horizon, out of sample. Each bar represents a specific equation as described in the legend, the x-axis depicts the lead in months of the corresponding equation, the y-axis gives the ratio of of mean absolute forecast error reduction. Reading example: At 15 months lead, a pooled equation with pseudo lagged forecast error reduces forecast errors to 82% of the original IMF forecast errors. A complete panel of countries may be found in the appendix.

South Korea, fare the worst. For Australia and Belgium for instance, the equations perform worse than the original IMF forecast over all horizons and settings. For France, the equations perform on average worse than the IMF forecast, only at the 10 months horizon a small forecast error reduction is possible. On the other hand, the equations perform reasonably well for smaller heavily globalised countries. Among the sample, Denmark, Finland, Luxembourg, the Netherlands, Sweden, New Zealand and the UK profit the most from the equations. Especially the equations using the pseudo lagged forecast errors as an additional regressor perform well, improving the forecasts at all horizons. For instance, UK these equation's forecast errors are about 0.9 times does of the original IMF forecasts at the 16 months horizon. At 14 months lead to 10 months lead they are down to 65 percent of the original forecast errors. Similarly for Luxembourg, where the IMF forecast error is reduced to 82% of its original value when using a pooled regression at a 15 months forecast horizon. Notable is the difference in forecast errors for the Euro Area 19, where the equations perform in a similar way than for the United Kingdom, reducing forecast errors from 0.85 to 0.5 times those originally published by the IMF.

As for the previous results, the country level shows a large difference between the lagged forecast error equation and those including early indicator variables. This is a promising sign, that the early indicator variables add valuable information to the forecasting process.

5 Real world application of the equations

During the creation of this work, it was decided to include preliminary studies on the underlying work in STATEC's "Note de Conjoncture". The "Note de Conjoncture" is a biannual economic publication, featuring the latest developments in the Luxembourgish economy as well as annual $t+0$ and $t+1$ forecasts for the most important macroeconomic aggregates. The target of this publication is the general public, hence the wording and explanations is adapted.

The original short study may be found online on the STATEC webpage (Note de Conjoncture 2 - 2021). It has been published in December 2020, using data up to beginning of October 2020. The external shock of the COVID19 recession is not ideal to test the results of this model, as neither its cause nor progress on its solution are tracked by economic indicators, especially not on an horizon of

16 to 10 months ahead. The panel equations are defined in percentage changes of real GDP forecast errors, while the COVID19 recession provokes an exogenous drop of demand which will be reversed as soon as the crisis is over. This means, that the shock is likely to imply a large rebound in 2021. This rebound however, can not be captured by the current set of variables in the model as the early indicator variables were still shocked by the pandemic in early October 2020. Still, this short study shows how the equations may be used to anticipate forecast revisions early on.

Figure 5: Original graph from Note de Conjoncture 2 - 2020

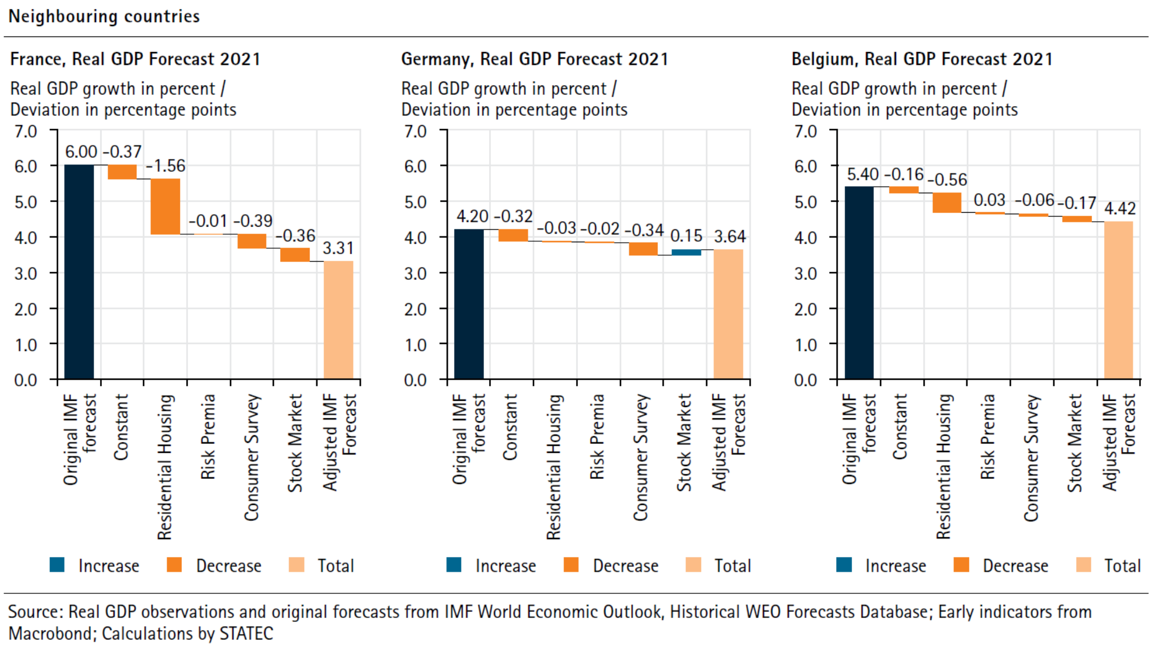

Figure 5: Original graph from Note de Conjoncture 2 - 2020, showing potential revisions of the respective IMF autumn 2020 forecasts for real GDP. The revisions are forecasted using the relationships between historical forecast errors and the early indicator variables. The forecast revisions are shown by contribution of early indicator.

Showing the direction of the probable revision, may help in updating forecasts and corresponding story telling. For instance figure 5 hint that french GDP might be more prone to downwards revision than the German one. A harsh drop in residential housing is the main contribution to this difference. Further, the stock market do not contribute into the same direction for all 3 sample countries, hence the stock market might be an inconclusive contribution to revision. If the model is run for several consecutive months, and the anticipated revision is strictly rising, at each iteration, the forecaster might be inclined to revise the next official update into the same direction. Further, looking into contributions of the single regressors gives an idea of the validity of the anticipated forecast revision produced by the panel equation.

6 Conclusion

The results are promising, there is room to improve year-ahead forecasts by using multi-country panel regressions and early information channels. This way, the IMF forecast may be tracked on a monthly basis. Within the IMF sample, mainly small open economies like Denmark, Finland, Luxembourg, the Netherlands or the United Kingdom profit from the panel equation. With the right specification of explanatory variables and a coherent panel of countries original IMF forecasts may be tracked and updated over the months.
Lagged forecast errors are found to improve the forecasts significantly. Currently, the equations are

relatively stable towards the coefficients. However, a major change in the IMF forecasting process might eradicate this feature. In the end the proposed model is a simple method to track year ahead international forecasts and detect changes in forecasts early on.

References

ARELLANO, Manuel; BOND, Stephen. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. The review of economic studies, 1991, 58. Jg., Nr. 2, S. 277-297.

ASTOLFI, Roberto, et al. The use of short-term indicators and survey data for predicting turning points in economic activity. 2016.

BANERJEE, Anindya; MARCELLINO, Massimiliano; MASTEN, Igor. Leading indicators for euro-area inflation and GDP growth. Oxford Bulletin of Economics and Statistics, 2005

CHEN, Sophia; RANCIERE, Mr Romain. Financial Information and Macroeconomic Forecasts. International Monetary Fund, 2017.

DEL NEGRO, Marco; SCHORFHEIDE, Frank. DSGE model-based forecasting. In: Handbook of economic forecasting. Elsevier, 2013. S. 57-140.

FIORAMANTI, Marco, et al. European Commission's Forecasts Accuracy Revisited: Statistical Properties and Possible Causes of Forecast Errors. 2016.

GENBERG, Hans; MARTINEZ, Andrew. On the accuracy and efficiency of IMF forecasts: A survey and some extensions. IEO Background Paper No. BP/14/04 (Washington: Independent Evaluation Office of the IMF), 2014.

GILCHRIST, Simon; SIM, Jae; ZAKRAJSEK, Egon. Uncertainty, financial frictions, and irreversible investment. Unpublished manuscript, 2013.

HAUSMAN, Jerry A. Specification tests in econometrics. Econometrica: Journal of the econometric society, 1978, S. 1251-1271.

MINCER, Jacob A.; ZARNOWITZ, Victor. The evaluation of economic forecasts. In: Economic forecasts and expectations: Analysis of forecasting behavior and performance. NBER, 1969. S. 3-46.

NICKELL, Stephen. Biases in dynamic models with fixed effects. Econometrica: Journal of the econometric society, 1981, S. 1417-1426.

PESARAN, M. Hashem. General diagnostic tests for cross section dependence in panels. 2004.

PESARAN, M. Hashem. Time series and panel data econometrics. Oxford University Press, 2015.

RATTO, Marco, et al. QUEST III: An estimated open-economy DSGE model of the euro area with fiscal and monetary policy. economic Modelling, 2009, 26. Jg., Nr. 1, S. 222-233.

WEO data base: https://www.imf.org/external/pubs/ft/weo/data/WEOhistorical.xlsx

TURNER, David. The use of models in producing OECD macroeconomic forecasts. 2016.

ZELLNER, Arnold. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. Journal of the American statistical Association, 1962, 57. Jg., Nr. 298, S. 348-368.

WOOLDRIDGE, Jeffrey M. Econometric analysis of cross section and panel data. MIT press, 2010.

Appendix A Appendix

Figure 6: Relative forecast error by country and forecast horizon.

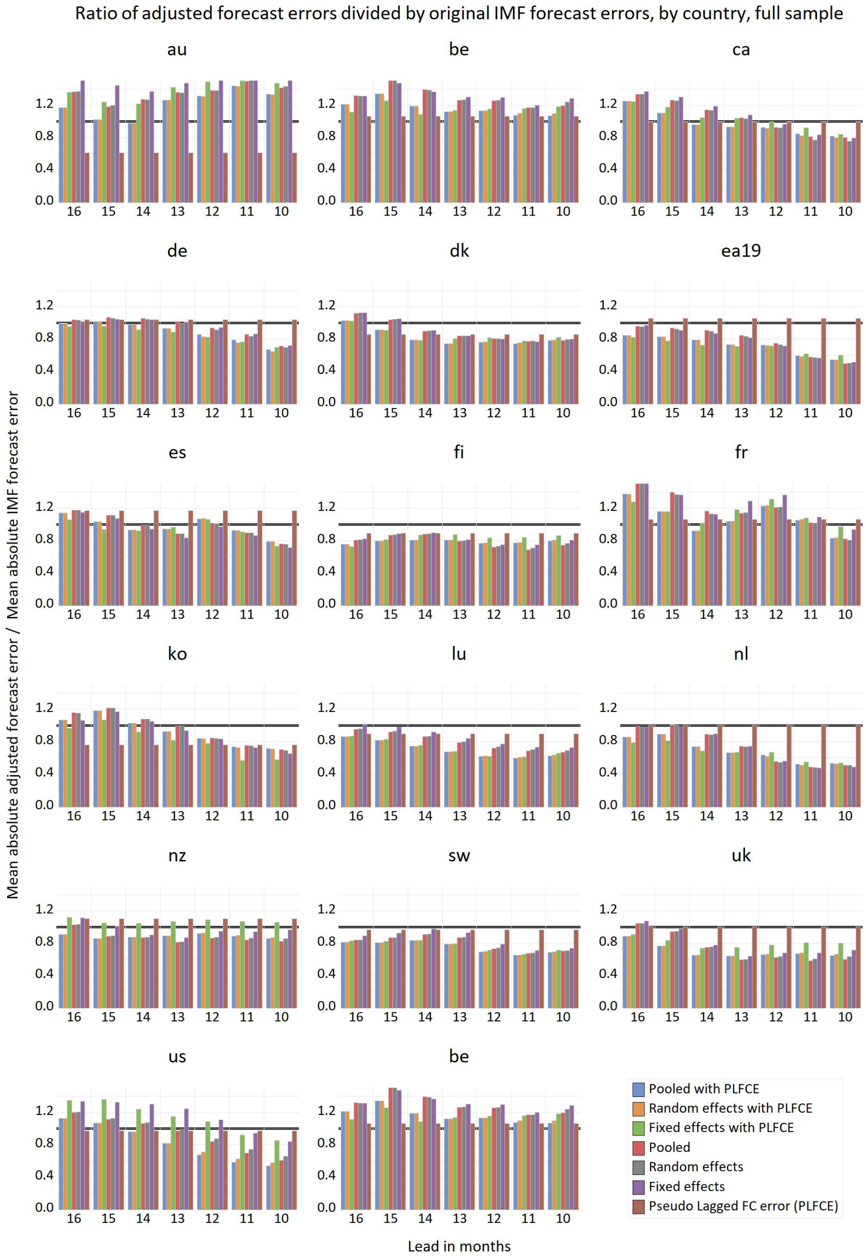

Figure 6: Bar plots are shown by country. Each bar represents a specific equation as described in the legend, the x-axis depicts the lead in months of the corresponding equation, the y-axis gives the ratio of of mean absolute forecast error reduction. Reading example: For the UK, at 15 months lead, a pooled equation with pseudo lagged forecast error reduces forecast errors to 76% of the original IMF forecast errors.

Figure 7: Mean out of sample forecast error reduction by year to forecast, lead in months and method.

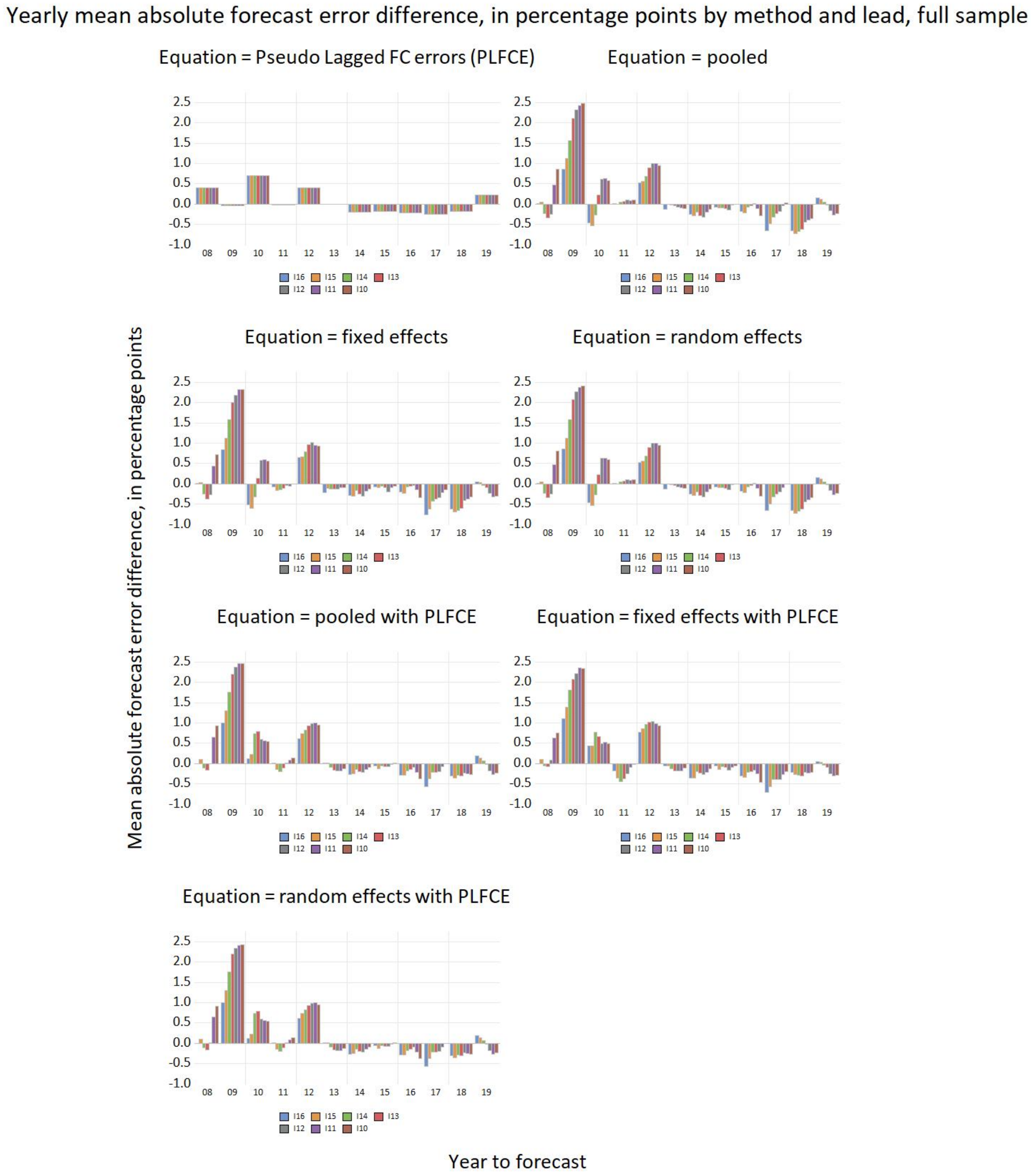

Table 4: Out of sample forecast error reduction with regard to historic IMF autumn forecast

Out of sample forecast error reduction with regard to the IMF autumn forecast, in percentage points: **European Sample**

Model:	16 months lead	15 months lead	14 months lead	13 months lead	12 months lead	11 months lead	10 months lead
Pool	-0.07	0.01	0.15	0.2*	0.24*	0.31**	0.36***
SUR	0.06	0.12	0.22**	0.26**	0.28**	0.33***	0.38***
Random	-0.08	0.01	0.14	0.19	0.23*	0.3**	0.34***
Fixed	-0.1	-0.02	0.11	0.16	0.2*	0.28**	0.32***
Pseudo Lagged FCError	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04	-0.04
Pool with pseudo lagged FCE	0.07	0.12	0.22*	0.22*	0.25*	0.33**	0.36***
Random with pseudo lagged FCE	0.14	0.19*	0.26**	0.29**	0.32**	0.35***	0.38***
SUR with pseudo lagged FCE	0.07	0.11	0.21*	0.22*	0.25*	0.32**	0.34***
Fixed with pseudo lagged FCE	0.07	0.09	0.18	0.19	0.2	0.29**	0.32***

Out of sample benchmarks using data from 1998 onwards, first out of sample realisation: 2008, last realisation: 2019
Pseudo lagged Forecast Error defined as *inyear forecast - year ahead forecast(-1)*
*** Significance at 0.01-level // ** Significance at 0.05-level // * Significance at 0.1-level

Table 4: This table compares the results of different panel equations with the original IMF forecasts. Mean absolute forecast errors of the original IMF forecasts are subtracted by the corresponding panel equation MAFEs. Each equation is run at different leads, from 16 months before the end of the year to 10 months before the end of the year. Hence, a 15 months lead corresponds to the moment of publication of the IMF forecasts. Due to restrictions of the SUR equation, the sample of countries is reduced to the European ones.
Reading example: A SUR equation with pseudo lagged forecast error as an additional regressor, targeting real GDP with 14 months lead, has on average a 0.18 percentage points lower forecast error than the original IMF forecast for the period 2008 to 2019.

Table 5: Variable description

MB identifier	Frequency	Description
resperm˙au	m	Australia’ OECD MEI’ Orders’ Construction’ Permits Issued’ Dwellings/Residential Buildings’ SA’ Index
gov10˙au	m	Australia’ Government Benchmarks’ Macrobond’ 10 Year’ Yield
gov2˙au	m	Australia’ Government Benchmarks’ Macrobond’ 2 Year’ Yield
survey˙au	m	Australia’ OECD MEI’ Consumer Opinion Surveys’ Confidence Indicators’ Composite Indicators’ OECD Indicator’ Normal=100’ SA’ Index
stockind˙au	m	Australia’ Equity Indices’ S&P/ASX’ 50’ Index’ Price Return’ Close’ AUD
resperm˙be	m	Belgium’ Construction Status’ Residential’ Belgium’ Buildings’ National Bank of Belgium (NBB)’ Permits
gov10˙be	m	Belgium’ Government Benchmarks’ Macrobond’ 10 Year’ Yield
gov2˙be	m	Belgium’ Government Benchmarks’ Macrobond’ 2 Year’ Yield
survey˙be	m	Belgium’ OECD MEI’ Consumer Opinion Surveys’ Confidence Indicators’ Composite Indicators’ OECD Indicator’ Normal=100’ SA’ Index
stockind˙be	m	Belgium’ Equity Indices’ Euronext Brussels’ BEL 20 Index’ Price Return’ Close’ EUR
resperm˙ca	m	Canada’ OECD MEI’ Orders’ Construction’ Permits Issued’ Dwellings/Residential Buildings’ SA’ Index
gov10˙ca	m	Canada’ Government Benchmarks’ 10 Year’ Yield
gov2˙ca	m	Canada’ Government Benchmarks’ 2 Year’ Yield
survey˙ca	m	Canada’ OECD MEI’ Consumer Opinion Surveys’ Confidence Indicators’ Composite Indicators’ OECD Indicator’ Normal=100’ SA’ Index
stockind˙ca	m	Canada’ Equity Indices’ S&P/TSX’ 60’ Index’ Price Return’ CAD
resperm˙de	m	Germany’ OECD MEI’ Orders’ Construction’ Permits Issued’ Dwellings/Residential Buildings’ SA’ Index
gov10˙de	m	Germany’ Government Benchmarks’ Macrobond’ 10 Year’ Yield
gov2˙de	m	Germany’ Government Benchmarks’ Macrobond’ 2 Year’ Yield
survey˙de	m	Germany’ OECD MEI’ Consumer Opinion Surveys’ Confidence Indicators’ Composite Indicators’ OECD Indicator’ Normal=100’ SA’ Index
stockind˙de	m	Germany’ Equity Indices’ Deutsche Boerse’ DAX’ 30 Index’ Price Return’ Close’ EUR
resperm˙dk	m	Denmark’ OECD MEI’ Orders’ Construction’ Permits Issued’ Dwellings/Residential Buildings’ SA
gov10˙dk	m	Denmark’ Government Benchmarks’ Macrobond’ 10 Year’ Yield
gov2˙dk	m	Denmark’ Government Benchmarks’ Macrobond’ 2 Year’ Yield
survey˙dk	m	Denmark’ OECD MEI’ Consumer Opinion Surveys’ Confidence Indicators’ Composite Indicators’ OECD Indicator’ Normal=100’ SA’ Index
stockind˙dk	m	Denmark’ Equity Indices’ Nasdaq OMX’ Benchmark’ OMXC20 Index’ Price Return’ Close’ DKK
resperm˙ea19	m	Euro Area 19’ OECD MEI’ Orders’ Construction’ Permits Issued’ Dwellings/Residential Buildings’ SA’ Index
gov10˙ea19	m	gov10˙ea19 = (gov10˙fr+gov10˙de+gov10˙es)/3
gov2˙ea19	m	gov2˙ea19 = (gov2˙fr+gov2˙de+gov2˙es)/3
survey˙ea19	m	Euro Area 19’ OECD MEI’ Consumer Opinion Surveys’ Confidence Indicators’ Composite Indicators’ OECD Indicator’ Normal=100’ SA’ Index
stockind˙ea19	m	Euro Area’ Equity Indices’ STOXX’ 50’ Index’ Price Return’ Close’ EUR
resperm˙es	m	Spain’ OECD MEI’ Orders’ Construction’ Permits Issued’ Dwellings/Residential Buildings’ SA’ Index
gov10˙es	m	Spain’ Government Benchmarks’ Macrobond’ 10 Year’ Yield
gov2˙es	m	Spain’ Government Benchmarks’ Macrobond’ 2 Year’ Yield
survey˙es	m	Spain’ OECD MEI’ Consumer Opinion Surveys’ Confidence Indicators’ Composite Indicators’ OECD Indicator’ Normal=100’ SA’ Index
stockind˙es	m	Spain’ Equity Indices’ Madrid Stock Exchange’ IBEX 35 Index’ Close’ EUR
resperm˙fi	m	Finland’ OECD MEI’ Orders’ Construction’ Permits Issued’ Dwellings/Residential Buildings’ SA’ Index
gov10˙fi	m	Finland’ Government Benchmarks’ Macrobond’ 10 Year’ Yield
gov2˙fi	m	Finland’ Government Benchmarks’ Macrobond’ 2 Year’ Yield
survey˙fi	m	Finland’ OECD MEI’ Consumer Opinion Surveys’ Confidence Indicators’ Composite Indicators’ OECD Indicator’ Normal=100’ SA’ Index
stockind˙fi	m	Finland’ Equity Indices’ Nasdaq OMX’ Benchmark’ OMX Helsinki 25 Index’ Price Return’ Close’ EUR
resperm˙fr	q	France’ OECD MEI’ Orders’ Construction’ Permits Issued’ Dwellings/Residential Buildings’ SA’ Index
gov10˙fr	m	France’ Government Benchmarks’ Macrobond’ 10 Year’ Yield
gov2˙fr	m	France’ Government Benchmarks’ Macrobond’ 2 Year’ Yield
survey˙fr	m	France’ OECD MEI’ Consumer Opinion Surveys’ Confidence Indicators’ Composite Indicators’ OECD Indicator’ Normal=100’ SA’ Index
stockind˙fr	m	France’ Equity Indices’ Euronext Paris’ CAC 40 Index’ Price Return’ Close’ EUR
resperm˙ko	m	South Korea’ OECD MEI’ Orders’ Construction’ Permits Issued’ Dwellings/Residential Buildings’ SA’ Index
gov10˙ko	m	South Korea’ Government Benchmarks’ Treasury Bonds’ 10 Year’ Yield
gov2˙ko	m	South Korea’ Government Benchmarks’ Treasury Bonds’ 3 Year’ Yield
survey˙ko	m	South Korea’ OECD MEI’ Consumer Opinion Surveys’ Confidence Indicators’ Composite Indicators’ OECD Indicator’ Normal=100’ SA’ Index
stockind˙ko	m	South Korea’ Equity Indices’ KOSPI’ KOSPI 50 Index’ Price Return’ Close’ KRW
resperm˙lu	m	resperm˙lu = resperm˙ea19 // Luxembourgish series very volatile
gov10˙lu	m	Luxembourg’ Government Benchmarks’ Central Bank of Luxembourg’ Bond Yields’ 10 Year
gov2˙lu	m	lu2ygov = nl2ygov // lu2ygov too short
survey˙lu	m	survey˙lu = survey˙ea19
stockind˙lu	m	Euro Area’ Equity Indices’ STOXX’ 50’ Index’ Price Return’ Close’ EUR
resperm˙nl	q	Netherlands’ Eurostat’ Building Permits’ Building Permits - Number of Dwellings’ Residential Buildings’ Except Residences for Communities ’Calendar Adjusted’ SA’ Index
gov10˙nl	m	Netherlands’ Government Benchmarks’ Macrobond’ 10 Year’ Yield
gov2˙nl	m	Netherlands’ Government Benchmarks’ Macrobond’ 2 Year’ Yield
survey˙nl	m	Netherlands’ OECD MEI’ Consumer Opinion Surveys’ Confidence Indicators’ Composite Indicators’ OECD Indicator’ Normal=100’ SA’ Index
stockind˙nl	m	Netherlands’ Equity Indices’ Euronext Amsterdam’ AEX All Share Index’ Price Return’ Close’ EUR
resperm˙nz	m	New Zealand’ OECD MEI’ Orders’ Construction’ Permits Issued’ Dwellings/Residential Buildings’ SA’ Index
gov10˙nz	m	New Zealand’ Government Benchmarks’ Reserve Bank of New Zealand’ 10 Year’ Yield
gov2˙nz	m	New Zealand’ Government Benchmarks’ Reserve Bank of New Zealand’ 2 Year’ Yield
survey˙nz	m	New Zealand’ OECD MEI’ Consumer Opinion Surveys’ Confidence Indicators’ Composite Indicators’ OECD Indicator’ Normal=100’ SA’ Index
stockind˙nz	m	New Zealand’ Equity Indices’ S&P/NZX’ 50 Index’ Total Return’ Close’ NZD
resperm˙sw	m	Sweden’ Eurostat’ Building Permits’ Building Permits - M2 of Useful Floor Area’ Residential Buildings (F˙CC11)’ 2015=100’ Calendar Adjusted’ SA’ Index
gov10˙sw	m	Sweden’ Government Benchmarks’ Macrobond’ 10 Year’ Yield
gov2˙sw	m	Sweden’ Government Benchmarks’ Macrobond’ 2 Year’ Yield
survey˙sw	m	Sweden’ OECD MEI’ Consumer Opinion Surveys’ Confidence Indicators’ Composite Indicators’ OECD Indicator’ Normal=100’ SA’ Index
stockind˙sw	m	Sweden’ Equity Indices’ Nasdaq OMX’ All-Share’ OMX Stockholm Index’ Price Return’ Close’ SEK
resperm˙uk	q	United Kingdom’ Eurostat’ Building Permits’ Building Permits - M2 of Useful Floor Area’ Residential Buildings’ Except Residences for Communities ’Calendar Adjusted’ SA’ Index
gov10˙uk	m	United Kingdom’ Government Benchmarks’ Macrobond’ 10 Year’ Yield
gov2˙uk	m	United Kingdom’ Government Benchmarks’ Macrobond’ 2 Year’ Yield
survey˙uk	m	United Kingdom’ OECD MEI’ Consumer Opinion Surveys’ Confidence Indicators’ Composite Indicators’ OECD Indicator’ Normal=100’ SA’ Index
stockind˙uk	m	United Kingdom’ Equity Indices’ FTSE’ 100’ Index’ Price Return’ Close’ GBP
resperm˙us	m	United States’ Leading Indicators’ Conference Board’ Business Cycle Indicators’ Fixed Capital Investment’ Building Permits for New Private Housing Units’ SA
gov10˙us	m	United States’ Government Benchmarks’ Federal Reserve’ 10 Year’ Yield
gov2˙us	m	United States’ Government Benchmarks’ Federal Reserve’ 2 Year’ Yield
survey˙us	m	United States’ OECD MEI’ Consumer Opinion Surveys’ Confidence Indicators’ Composite Indicators’ OECD Indicator’ Normal=100’ SA’ Index
stockind˙us	m	United States’ Equity Indices’ S&P’ 500’ Index’ Price Return’ Close’ USD

Figure 8: Rolling regression estimates of a pooled model

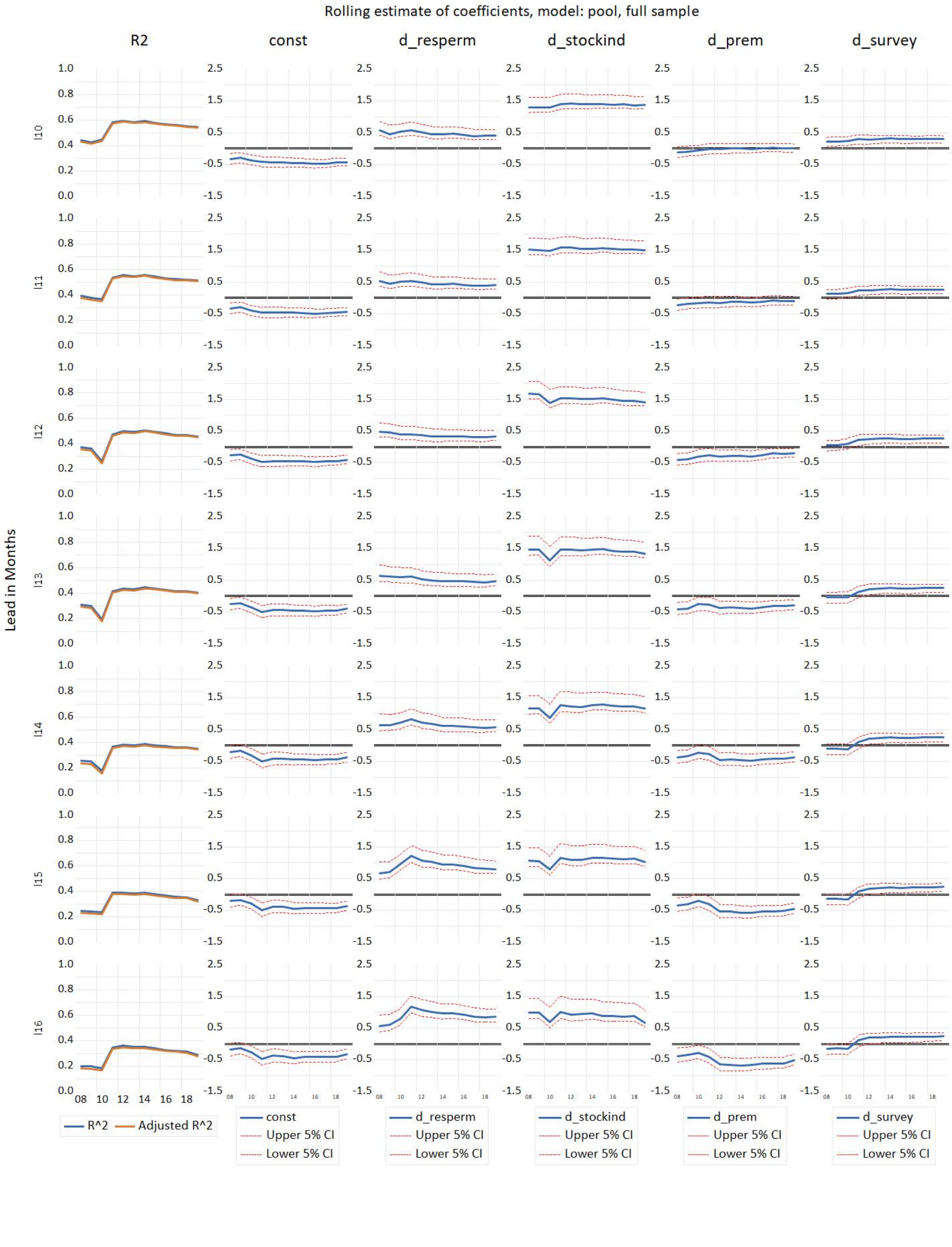

Figure 9: Rolling regression estimates of a fixed effects panel

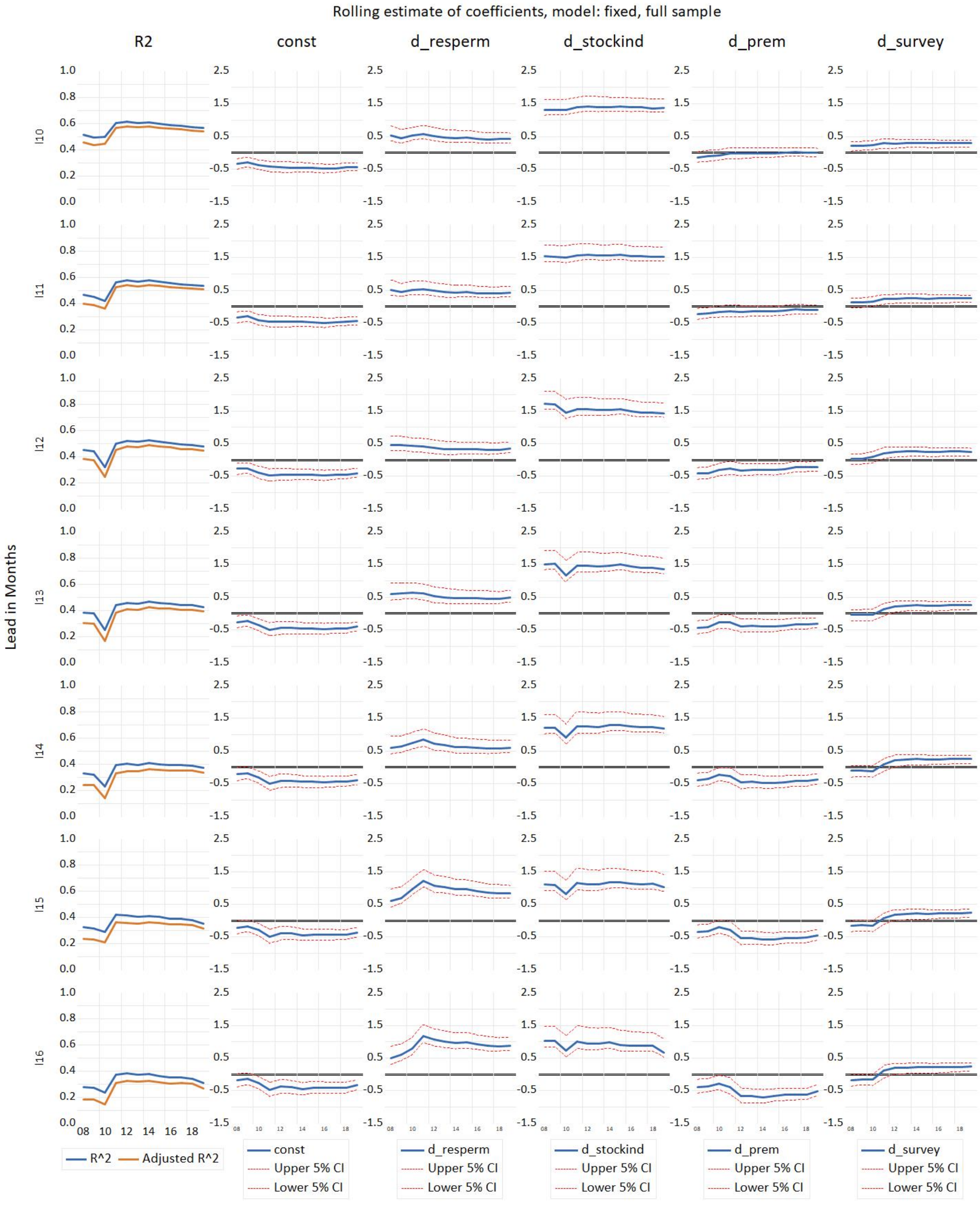

Figure 10: Rolling regression estimates of a random effects model

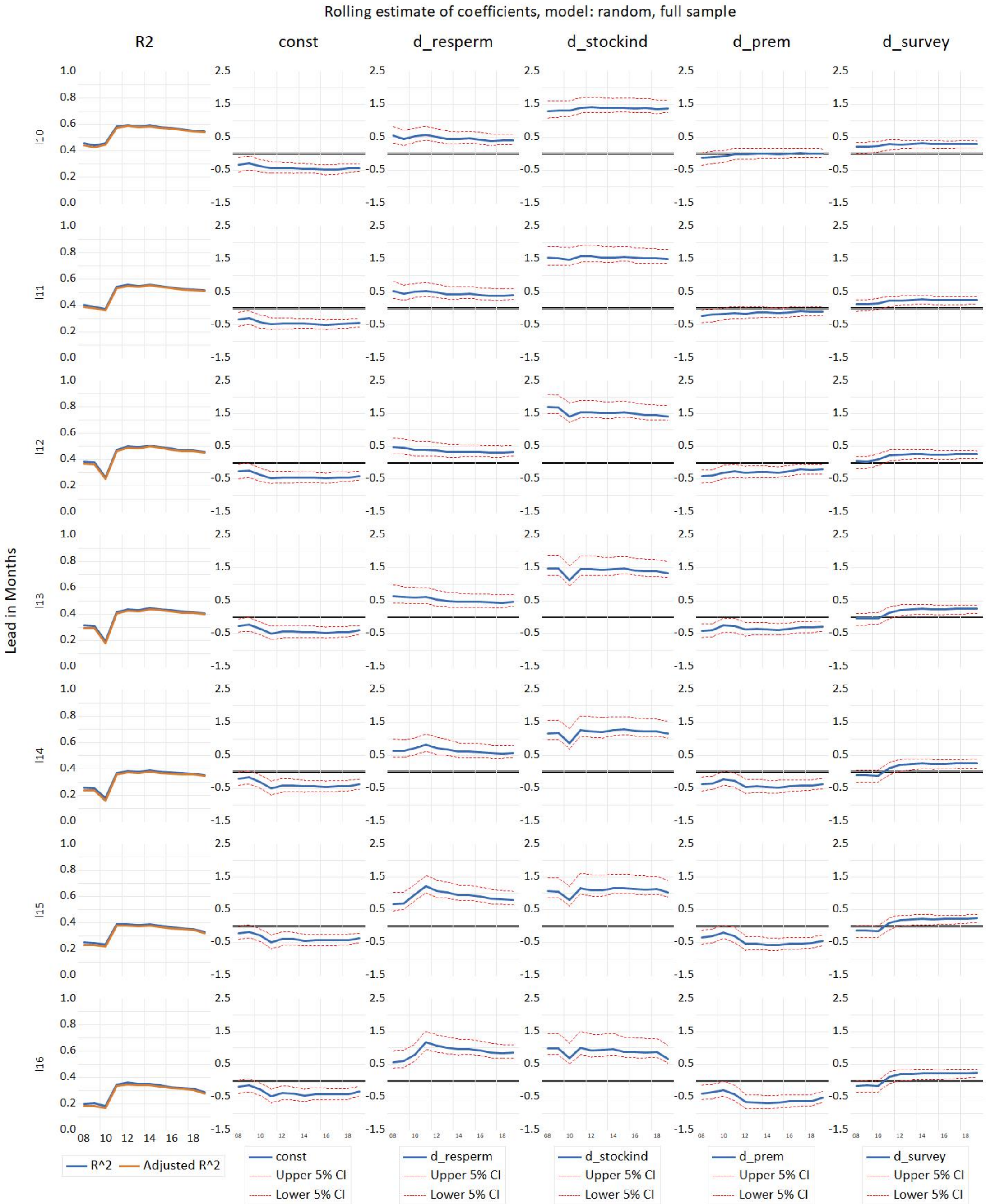

Figure 11: Rolling regression estimates of a SUR model, European sample

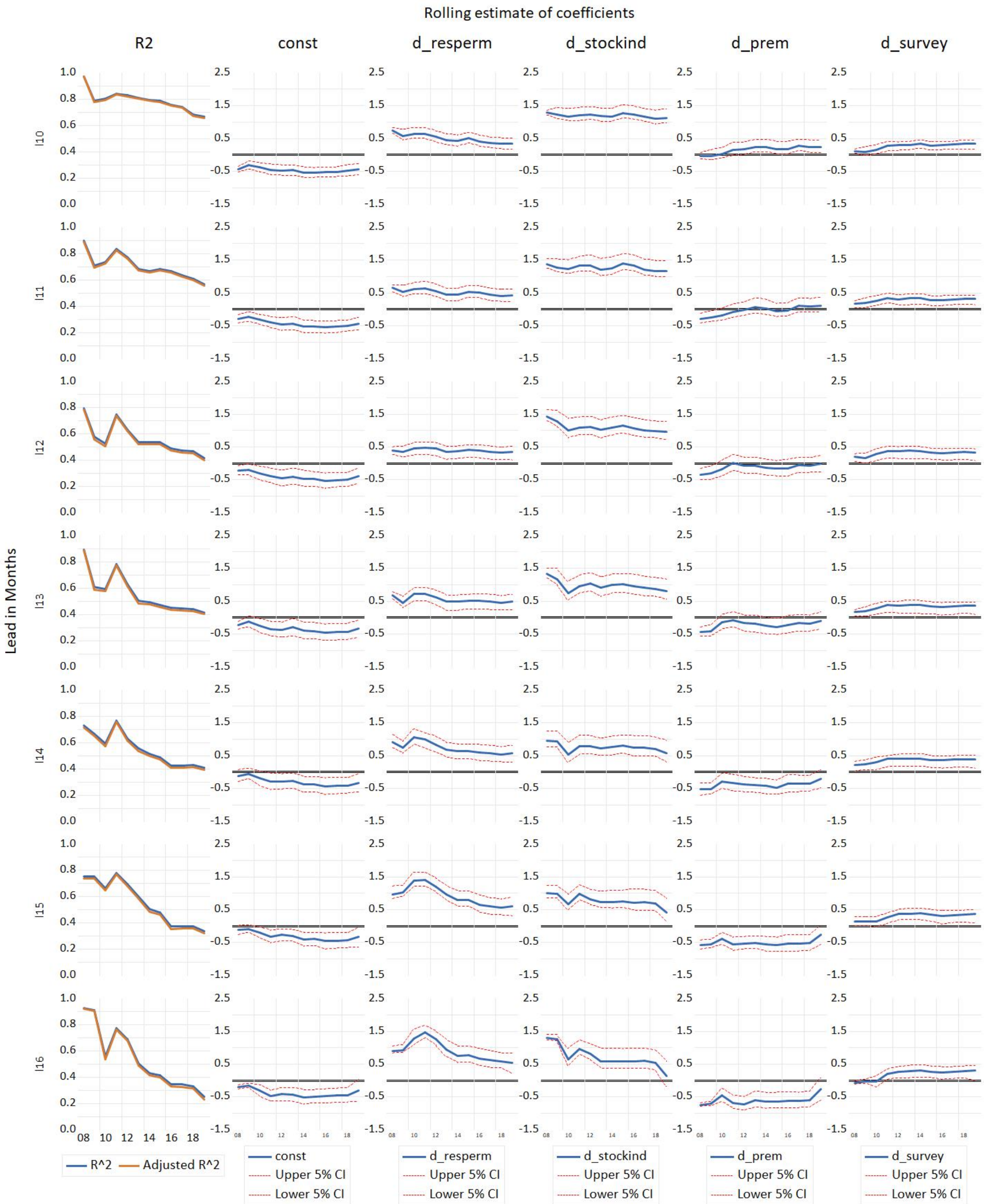

Figure 12: Rolling regression estimates of a pooled model with a pseudo lag of the forecast error as an additional regressor

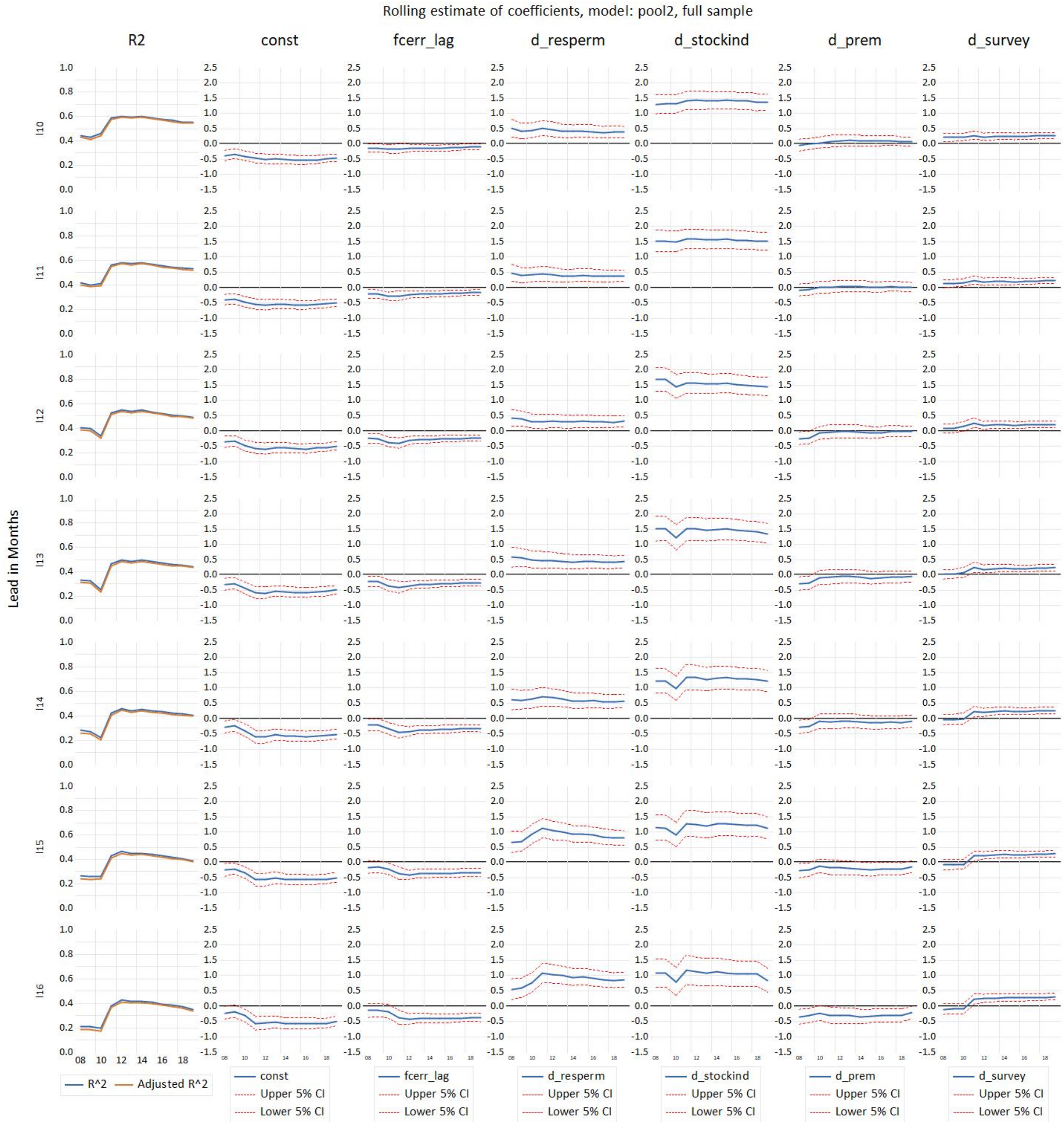

Figure 13: Rolling regression estimates of a fixed effects panel model with a pseudo lag of the forecast error as an additional regressor

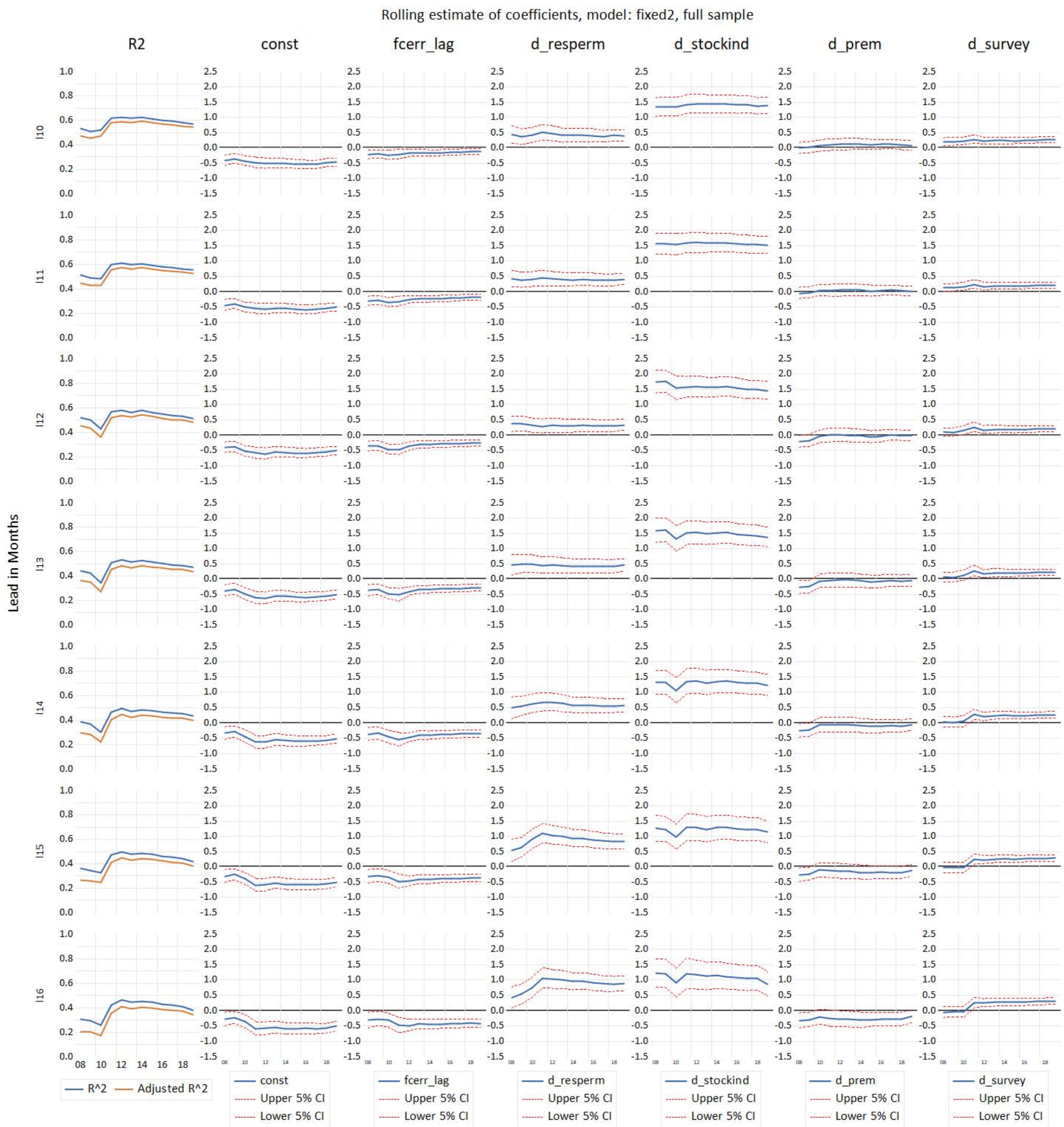

Figure 14: Rolling regression estimates of a random effects panel model with a pseudo lag of the forecast error as an additional regressor

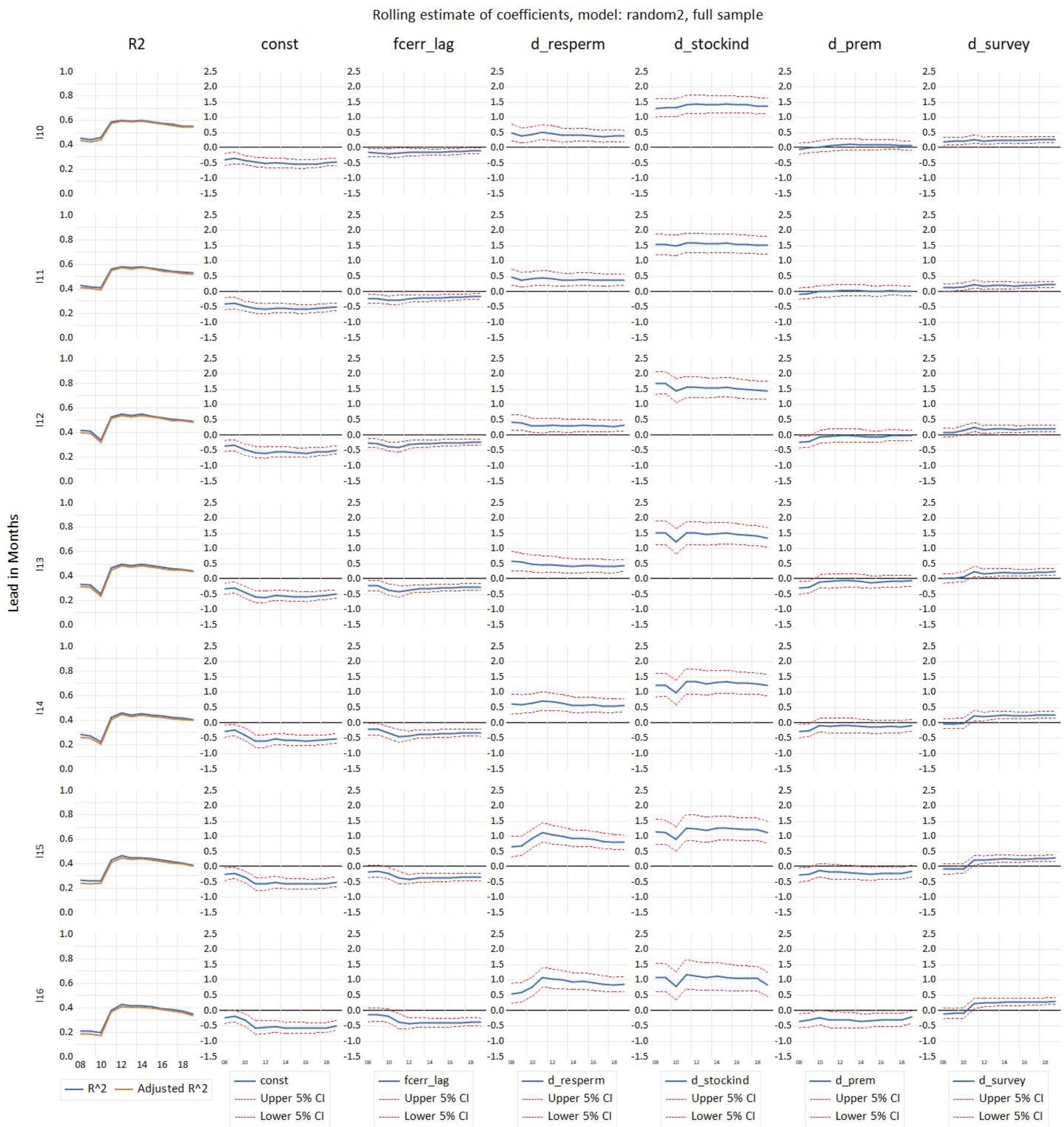

Figure 15: Rolling regression estimates of a SUR model with a pseudo lag of the forecast error as an additional regressor, European sample