Why can't professional macroeconomic forecasters predict recessions?

Víctor López Pérez[1]

23 May 2019

Abstract

The professional forecasters' inability to anticipate macroeconomic recessions is well documented. The literature has found that aggregate or consensus forecasts are too optimistic before downturns and too pessimistic before recoveries. This paper explores whether this result also holds with individual data or is the result of an aggregation bias. Using a Spanish survey of professional forecasters conducted by Funcas, I find that forecasters are indeed too optimistic before recessions for two reasons. First, strong herding behaviour around the consensus forecast prevents those forecasters perceiving the early signs of a recession from adjusting their expectations as much as needed to predict it. And second, some forecasters put too much weight on the most recent developments when producing their forecasts: better-than-expected data makes some forecasters to revise their forecasts upwards too much. These revisions raise the consensus forecast and trigger herding by other forecasters, who also revise up their forecasts. Both factors lead to subsequent negative forecast errors, especially when a recession occurs. Consequently, professional forecasters could improve their forecasting performance by placing less weight on indicators from the recent past and by avoiding inefficient herding.

Keywords: professional forecasters, Funcas, forecast errors, forecast revisions, expected GDP growth.

JEL classification: D84, E37.

[1] Author's affiliation: Universidad Politécnica de Cartagena, Facultad de Ciencias de la Empresa, Department of Economics, c/ Real 3, 30201, Cartagena, Spain. e-mail: victor.lopez@upct.es. Phone number: (+34) 868071259. Fax number: (+34) 968325781. I am grateful to Funcas, and in particular to María Jesús Fernández Sánchez, for providing the dataset and for very helpful technical assistance. I thank David Gabauer, Ayaz Zeynalov and participants to the 9th International Conference in Economics (Econworld 2019, Seville, Spain) for interesting comments and discussions. I am also grateful to Susana Tena Nebot for excellent research assistance.

1. Introduction

Professional macroeconomic forecasters are experts that predict where the economy of a country, region or the world is going. They are typically econometricians or statisticians working for private companies or public institutions.[2] Their views are highly appreciated in financial and policy circles as they provide valuable insights on future economic developments. For this reason, the US Federal Reserve, the European Central Bank, the Bank of Japan and the Bank of England, among others, conduct surveys of professional forecasters regularly, typically with a quarterly frequency. And the results of these surveys are used as an input to inform decisions by policy-makers.

But how accurate these professional forecasters are? There is a developing literature on forecasting evaluation and performance applied to the predictions of professional forecasters with mixed results. Some authors have found that these predictions are irrational (Ager *et al*., 2009, Rossi and Sekhposyan, 2016, An *et al.* (2018), 2018, Gelain *et al*., 2019), biased (Bonham and Cohen, 2001, Harvey and Newbold, 2003, Garcia and Manzanares, 2007, Capistran and Timmermann, 2009, Andrade and Le Bihan, 2013, Best and Kapinos, 2018, Ramos-Herrera and Sosvilla-Rivero, 2018, Rossi and Sekhposyan, 2018) or inferior to other alternatives (Wieland and Wolters, 2011, Baghestani, 2019, Davig and Hall, 2019, Galbraith and van Norden, 2019). Others have argued that professional forecasts are rational (Deschamps and Ioannidis, 2013, Wang and Lee, 2014, Coibion and Gorodnichenko, 2015, El-Shagi, 2018), unbiased (Croushore, 2010, Frenkel *et al*., 2011) and perform well compared to other alternatives (Ang *et al.*, 2007, Rubaszek and Skrzypcznski, 2008, Grothe and Meyler, 2018, Zhang, 2018, Gelfer, 2019). Finally, some papers explored performance at the individual level and found significant heterogeneity, with some forecasters performing much better than others (Dovern and Weisser, 2011, Poncela *et al*., 2011, Conflitti *et al*., 2015, López Pérez, 2016a, Diebold and Shin, 2018).

A recent paper by Dovern and Jannsen (2017) shed new light on this debate. They used Consensus Economics' aggregate forecasts of annual real GDP growth for 19 developed countries over the sample 1990-2013 and found that their bias depends on the phase of the economic cycle: forecasts are unbiased during expansions, way too optimistic before recessions and mildly pessimistic before recoveries. In other words, forecasters are unable to predict the turning points of the economic cycle. They cannot predict recessions.[3]

I build upon Dovern and Jannsen' paper along three dimensions. First, I test whether their results with aggregate forecasts also apply to individual forecasts or are just the result of an aggregation bias (Keane and Runkle, 1990, Bonham and Cohen, 2001). If, for instance, at the end of an economic expansion, half of the forecasters predict the start of a recession correctly while the other half are inattentive and forecast that the expansion would go on, the aggregate forecast would be too optimistic when the recession hits. But this does not mean that professional forecasters cannot predict recessions, simply that there is significant heterogeneity among them. In the appendix of their paper, Dovern and Jannsen used individual forecasts from the US SPF to investigate whether their results hold at the individual level. However, as I showed in a previous paper (López Pérez, 2016b), surveys like the US SPF or the European Central Bank's SPF may be subject to important sample-composition effects at the turns of the business cycle, and Dovern and Jannsen did not control for them. In this paper I use a different database of individual forecasts that is not affected by sample-composition effects at the turns of the business cycle.

Second, I inform about the reasons why professional forecasters seem unable to predict recessions, a question that has even made it to the financial news media recently.[4] Some authors have suggested that forecasters are

[2] For an incomplete list of the professional forecasters surveyed by the European Central Bank see https://www.ecb.europa.eu/stats/ecb_surveys/survey_of_professional_forecasters/html/index.en.html. For a list of those surveyed by the Federal Reserve Bank of Philadelphia see https://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters/2018/survq318.

[3] This point was also made by Best and Kapinos (2018), for the aggregate forecast of GDP growth obtained from the US Survey of Professional Forecasters (SPF). An *et al.* (2018) found that one-year-ahead aggregate forecasts from the Consensus Economics survey accurately predicted 5 out of 153 recessions only. Rudebusch and Williams (2009) concluded that "professional forecasters are worse at predicting recessions a few quarters ahead than a simple real time forecasting model that is based on the yield spread".

[4] See the article "Why are economists so bad at forecasting recessions?" by S. Kennedy and P. Coy, published on Bloomberg Businessweek on 28 March 2019.

rational but inattentive and thereby may fail to take into account useful available data that would help them forecast recessions (Nordhaus, 1987, Mankiw and Reis, 2002, Reis, 2006). Others have proposed that forecasters are rational but receive noisy information, a feature that may make them too cautious when revising their forecasts (Woodford, 2002, Sims, 2003, Coibion and Gorodnichenko, 2012, Dovern, *et al*., 2015). Finally, some have argued that forecasters are irrational because they are too backward-looking (Gelain *et al.*, 2019), which may lead to excessive optimism during good times. In this paper I present empirical models that identify the main drivers of forecast errors and forecast revisions, which provide additional insights on the channels that are preventing professional forecasters from predicting recessions.

And third, I extend the analysis to the subcomponents of GDP: consumption, investment, government expenditure, exports and imports. The reason is that professional forecasters may be unable to predict sudden shifts in the most volatile components of GDP, like investment and exports, but may be more successful with the most persistent ones, like consumption and imports. If this were the case, professional forecasters would be able to predict recessions better if they improved the performance of their models for a few clearly identified components of GDP.

As a preview, the results of the paper are consistent with the hypothesis that individual forecasters are irrational in three ways. First, they seem to be partly backward-looking, a feature that makes them too optimistic before recessions. This finding applies to almost all subcomponents of GDP, no matter how volatile they are, suggesting that something more fundamental is at the root of the inability to predict the turning points of the cycle. Second, they seem to herd around the consensus or average forecast, which reduces the attractiveness of large forecast revisions when the first signs of a recession appear. And third, forecast revisions are autocorrelated at the individual level, a result that is at odds with the hypothesis of rational expectations.

The rest of the paper is structured as follows. Section 2 describes the dataset used. Section 3 presents the estimations of empirical models of forecast errors based on Dovern and Jannsen (2017). Section 4 shows the estimations of the model of forecast revisions. Section 5 concludes and outlines directions for further research.

2. The data

2.1 The Funcas database

The database used in this paper is maintained by Funcas, a non-profit foundation created by the Spanish Federation of Savings Banks. Funcas conducts a survey six times a year to ask professional forecasters located in Spain about future expected macroeconomic variables of the Spanish economy. The respondents are mostly financial institutions but also universities and non-financial corporations. Annex I lists the names of all the contributors to this survey and the date when they first replied.

The variables surveyed by Funcas and used in this paper are the expected annual growth rates of real GDP, private consumption, government expenditure, investment, exports and imports. These expectations are for the current and the next calendar year. Therefore, they are fixed-event forecasts and the forecast horizon decreases during the calendar year. These forecasts are available since May 1999 with the exception of the expected growth rate of government expenditure, whose inclusion in the survey questionnaire was discontinued between June 1999 and October 2005.[5]

Additionally, Funcas surveys forecasts of quarter-on-quarter real GDP growth, quarter-on-quarter growth rates of private consumption and quarter-on-quarter CPI inflation.[6] These forecasts are not published and Funcas kindly made them available upon request under the condition that the individual data cannot be identifiable.

[5] The Funcas survey also includes expectations of other variables not used in this paper: the annual growth rates of investments in machinery and capital goods, construction, national demand, the consumer price index (CPI), the core CPI, labour costs and employment, the average unemployment rate, the average current-account balance and the average public deficit.

[6] Quarter-on-quarter forecasts have replaced year-on-year forecasts for each quarter since September 2010.

The Funcas survey has advantages and disadvantages with respect to other surveys of professional forecasters. Its main disadvantage versus the Consensus Economics survey is that its forecasts are for one country only.[7] In this regard, it is similar to other widely used surveys like the US SPF or the ECB's SPF. Its main disadvantage versus the SPFs is that the number of forecasters is relatively small: the maximum number of forecasters per survey round is 20, in line with the Consensus Economics survey.

On the advantages, the Funcas survey is free of charge, unlike Consensus Economics'; it surveys much more variables than the ECB's SPF, including the components of GDP, which are one of the focuses of this paper; the survey is conducted six times a year, versus four times a year in the US SPF and the ECB's SPF; and it has a negligible number of missing observations for active forecasters, unlike the US SPF or the ECB's SPF, a feature that allows the analysis of individual survey data without the need to control for sample-composition effects at the turns of the business cycle.

Regarding the length of the forecast horizon, professional forecasters are requested by Funcas to submit expectations for the current and the following calendar years in survey rounds from March to December (see the months when the survey has been conducted in Annex II). In the surveys taking place in January or February, expectations refer to the previous and the current calendar year. There is one exception: in February 2012, forecasters submitted forecasts for 2012 and 2013. Therefore, if the horizon length of a forecast for year *t* submitted in February of year *t+1* is normalised to zero, the range of forecast horizon lengths goes from 24 months (forecast for 2013 submitted in February 2012) to 0 months (forecast for year *t* submitted in February of year *t+1*).

2.2 Realisations

Forecasts are subtracted from realisations to compute forecast errors. The realisations are obtained from the OECD real-time database, following Dovern and Jannsen. Real-time data is used because forecasters cannot predict future methodological changes to the way target variables are calculated (Harvey and Newbold, 2003). More precisely, forecasts for year *t* are compared to the realisation from the vintage of data available in March of year *t+1*, when fourth-quarter data from the previous year is available for the first time.[8]

2.3 Expansion, recession and recovery years

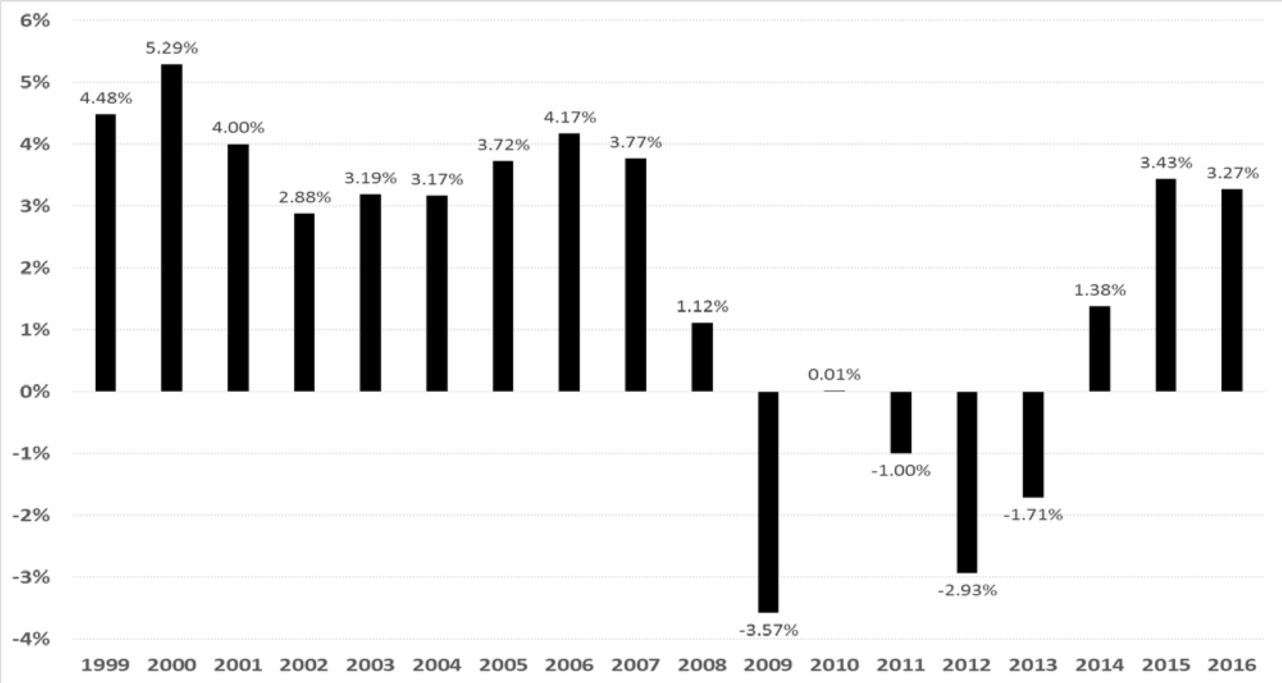
One of the contributions by Dovern and Jannsen (2017) is to show that the sign of the forecast errors depends on the phase of the business cycle: close to zero during expansions, negative in recessions and positive in recoveries. To that end, they defined year *t* as a recession year if real GDP in *t* is lower than in *t-1*. Year *t* is defined as a recovery year if real GDP in *t* is higher than in *t-1* and the latter is a recession year. Finally, year *t* is defined as an expansion year if real GDP in *t* is higher than in *t-1* and the latter is not a recession year.

In this paper I follow Dovern and Jannsen and use the most recent vintage available of GDP data to identify expansions, recessions and recoveries in the Spanish economy. Revised data is obtained from the Federal Reserve Bank of St. Louis (FRED database). Figure 1 shows Spanish real GDP growth rates. According to the metric used, there are four recession years in the sample (2009, 2011, 2012 and 2013), two recovery years (2010 and 2014) and the rest are expansion years.

[7] It may be argued that another disadvantage of the Funcas survey versus the Consensus Economics survey is that the latter is conducted monthly. However, ECB (2019) shows that 74% (64%) of a sample of professional forecasters update their short-term (medium-term) forecasts of GDP growth once a quarter, while only 19% (11%) update once a month.

[8] There were a few years when fourth-quarter data of real GDP was available in February according to the OECD real-time database. This data, however, is disregarded for the calculation of the realisations because it was inconsistent with the first vintage of data for the previous full calendar year published by the Bank of Spain in the March edition of its Economic Bulletin. The data available in March, according to the OECD database, is always consistent with the data published by the Bank of Spain.

Figure 1. Real GDP growth rates in Spain (revised data)



Source: Own calculations based on data from the Federal Reserve Bank of Saint Louis.

2.4 Forecast errors

As said above, forecast errors are defined as follows:

$$FE_{i,x,t,h} = R_{x,t} - F_{i,x,t,h} \tag{1}$$

where $FE_{i,x,t,h}$ is the forecast error made by forecaster i when predicting variable x for year t with a forecast-horizon length of h months, $R_{x,t}$ is the realisation of variable x in year t, and $F_{i,x,t,h}$ is the forecast submitted by forecaster i of variable x for year t with a forecast horizon length of h months. Average forecast errors by forecast horizon can be obtained by running the following regression:

$$FE_{i,x,t,h} = \Sigma_h D_h \alpha_h + \varepsilon_t \qquad for\ h=0,1,2,..,24. \tag{2}$$

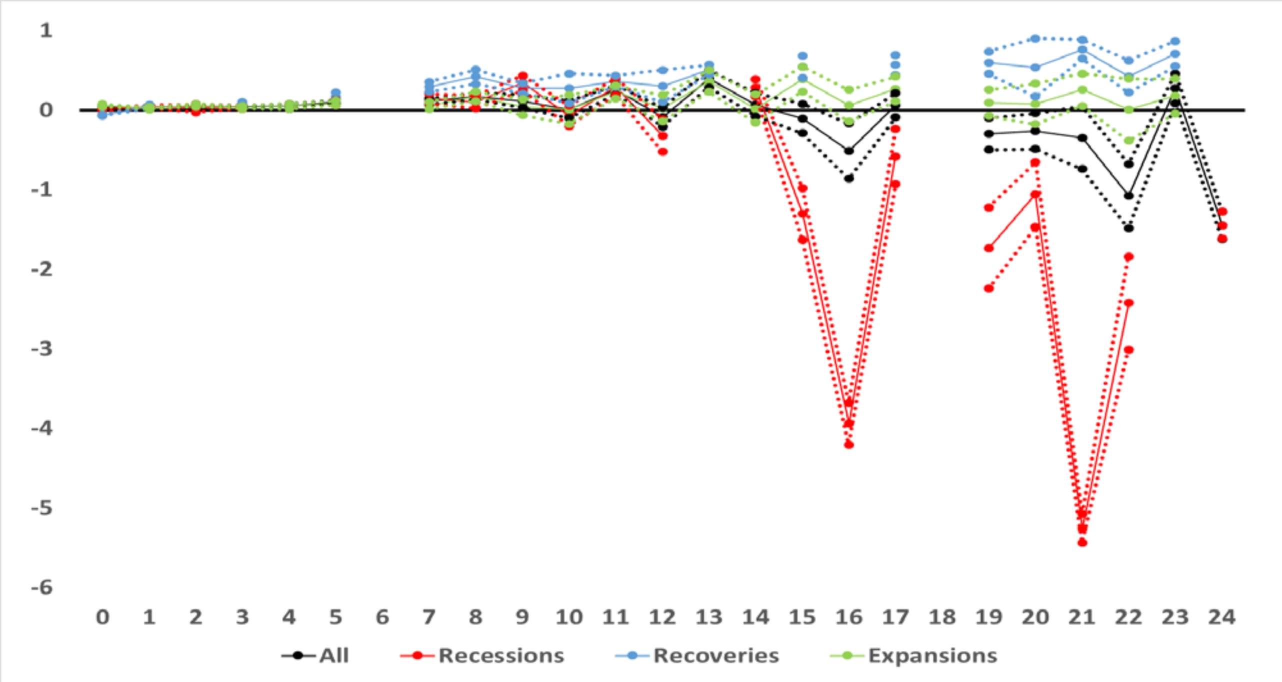
where D_h is a dummy variable that takes a value of 1 when the forecast horizon length is h and a value of 0 otherwise, α_h is a horizon-specific constant term and ε_t is a zero-mean disturbance. The model is estimated by pooled OLS (ordinary least squares) over a sample from May 1999 to January 2017. The standard errors of the α_h have been computed with bootstrap[9] due to the likely presence of cross-correlation and heteroskedasticity in the error term: cross-correlation may arise because of the fixed-event nature of the forecasts, with several forecasts for the same variable submitted in different months, and heteroskedasticity may arise because the variance of the error term is likely to increase with the length of the forecast horizon (Ager *et al.*, 2009).

Figure 2 shows the regression results for GDP-growth forecasts. The black solid line depicts α_h for each h. The black dotted lines are the bounds of their 95% confidence intervals. Average forecast errors are very close to zero for the shortest forecast horizons, but they became slightly negative for the longest ones. When the same regression is run for forecasts whose target year is a recession year, I obtained the estimates shown in red: much larger negative forecast errors for horizons longer than 14 months, a result that is consistent with the asymmetric forecast errors documented by Harvey and Newbold (2003) and Galbraith and van Norden (2019). For recovery years, the regression results are shown in blue: significantly positive forecast errors for horizons longer than 6 months. Finally, for expansion years, average forecast errors (in green) are statistically very close to zero for all horizons.

[9] Across 3000 replications with 3054 observations each.

These results are consistent with those reported by Dovern and Jannsen using the Consensus Economics survey: professional forecasters are too optimistic before recessions, too pessimistic before recoveries and unbiased during expansions. As the average size of the forecast errors before recessions is much larger than before recoveries, forecasts appear to be biased on average, especially for the longest horizons.

Figure 2. Average forecast errors for real GDP growth by horizon lengths and phase of the business cycle.



Annex III shows the same statistics for the subcomponents of GDP: private consumption, investment, government expenditure, exports and imports. The pattern of forecast errors for consumption, investment, exports and imports across the business cycle is very similar to that of GDP's, with negative forecast errors during recession years and positive forecast errors during recoveries. Having said this, some interesting results arise. Forecasts of growth rates of exports and imports are too high not only during recessions but also during expansions. This result is at odds with Dovern and Jannsen's conclusion that professional forecasts are unbiased conditional on the economy staying in an expansion.

Similarly, growth forecasts of government expenditure appear to be too low not only during recessions, when countercyclical fiscal policies are likely to be implemented, but also during expansions. This is another finding that is not consistent with the hypothesis that professional forecasts are unbiased during expansions. Instead, forecasts of GDP growth during expansions happen to be unbiased because the negative biases in exports and imports and the positive bias in government expenditure broadly cancel out.

Finally, a caveat must be made about the differences between the two recoveries in the sample, 2010 and 2014. As shown in Figure 1, GDP growth in 2010 was only 0.01% in the revised dataset and it was negative in the real-time first estimate published in March 2011 (-0.15%). Therefore, 2010 looks rather different in many ways when compared to a standard recovery year like 2014, whose GDP growth rate was well above zero (1.38%). For example, the average forecast error of the growth rate of investment in 2010 is negative for forecast horizon lengths longer than 7 months. The heterogeneity across the two recoveries increases the uncertainty surrounding the estimates for recoveries presented below and thereby they should be taken with caution.

Table 1. Estimation results of the empirical model of GDP-growth forecast errors.

	M1	M2	M3	M4	M5	M6	M7	M8
β_1 (recession)		-0.876	-0.847	-2.613	0.436	-0.147	-0.603	-0.703
		(0.00)	(0.00)	(0.00)	(0.00)	(0.11)	(0.00)	(0.00)
β_2 (recovery)			0.131	-1.626	-0.007	-0.599	-0.323	-0.469
			(0.00)	(0.00)	(0.92)	(0.00)	(0.00)	(0.00)
δ_1 (recession started)				2.125	-0.013	-0.372	0.256	-0.501
				(0.00)	(0.86)	(0.00)	(0.01)	(0.00)
δ_2 (recovery started)				1.905	-0.287	-0.917	-0.358	-0.892
				(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ_1					-0.012	-0.016	-0.014	0.013
					(0.00)	(0.00)	(0.00)	(0.05)
γ_1 squared								-0.001
								(0.00)
γ_2 (recession)					-0.194	-0.174	-0.141	-0.073
					(0.00)	(0.00)	(0.00)	(0.01)
γ_2 squared								-0.003
								(0.01)
γ_3 (recovery)					-0.078	-0.076	-0.059	-0.190
					(0.00)	(0.00)	(0.00)	(0.00)
γ_3 squared								0.005
								(0.00)
γ_4 (recession started)					0.112	0.114	0.086	0.157
					(0.00)	(0.00)	(0.00)	(0.00)
γ_4 squared								-0.002
								(0.19)
γ_5 (recovery started)					0.133	0.101	0.061	0.181
					(0.00)	(0.00)	(0.00)	(0.00)
γ_5 squared								-0.004
								(0.00)
λ_1						-0.276		-0.205
						(0.00)		(0.00)
λ_2						0.172		0.109
						(0.00)		(0.04)
λ_3						-0.028		-0.138
						(0.56)		(0.00)
λ_4						-0.362		-0.414
						(0.00)		(0.00)
λ_5						-0.392		-0.393
						(0.00)		(0.00)
λ_6						-0.280		-0.174
						(0.00)		(0.00)
λ_7						-0.084		-0.047
						(0.06)		(0.30)
λ_8						0.210		0.161
						(0.00)		(0.00)
$\Sigma\lambda_i$ (expansion)							-0.750	
							(0.00)	
$\Sigma\lambda_i$ (recession)							-1.666	
							(0.00)	
$\Sigma\lambda_i$ (recovery)							0.429	
							(0.01)	
α	-0.041							
	(0.02)							
Observations	3054	3054	3054	3037	3037	2808	2808	2808
Adj. R^2	0.00	0.16	0.16	0.39	0.53	0.62	0.67	0.65

Notes: P-values in parenthesis. Models M2 to M8 include individual dummy variables.

3. The empirical model of the forecast errors

3.1. The Dovern and Jannsen (2017) model

Dovern and Jannsen suggested to move past the simple model presented in the previous section and estimate a model of the forecast errors with individual effects and interactions between the phase of the business cycle and the length of the forecast horizon. Therefore, I estimate the following model:

$$FE_{i,GDP,t,h} = \Sigma_i \, \alpha_i \, D_i + (\beta_1 + \gamma_2 \, h) \, D_{recession} + (\beta_2 + \gamma_3 \, h) \, D_{recovery} + (\delta_1 + \gamma_4 \, h) \, D_{recession\ started} + \\ + (\delta_2 + \gamma_5 \, h) \, D_{recovery\ started} + \gamma_1 \, h + \varepsilon_t \qquad for\ i=1,2,3,\ldots 23 \qquad (3)$$

where $FE_{i,GDP,t,h}$ is the forecast error made by forecaster *i* when predicting GDP growth for year *t* with a forecast-horizon length of *h* months, D_i is an individual dummy variable that takes a value of 1 for forecaster *i* and 0 otherwise, *h* is the length of the forecast horizon in months, $D_{recession}$ and $D_{recovery}$ are dummy variables that take a value of 1 if the target year of the forecast is a recession or a recovery year respectively, $D_{recession\ started}$ and $D_{recovery\ started}$ are dummy variables that take a value of 1 if the forecast was submitted during a recession or a recovery year respectively, and ε_t is a random disturbance with zero mean. $\alpha_1, \ldots, \alpha_{23}$, β_1, β_2, δ_1, δ_2, γ_1, γ_2, γ_3, γ_4 and γ_5 are the model parameters, estimated by pooled OLS with the same sample used in subsection 2.4. Table 1 shows the estimated coefficients and their p-values based on bootstrap standard errors.

Model 1 (M1) is the result of imposing the restriction that all the parameters in (3) are equal to zero with the exception of a constant term. This constant is negative and statistically significant, which means that professional forecasters were 4 basis points too optimistic on average when they predicted GDP growth. In model 2 (M2) the constant is replaced with 23 individual effects (α_i) to account for *micro-heterogeneity* in pooled estimators (Bonham and Cohen, 2001) and β_1 is no longer assumed to be zero. The estimated β_1 is negative and statistically significant, which means that forecasts of GDP growth are almost 1 percentage point too high on average when the target year is a recession year.

In model 3 (M3), β_2 is also estimated. It is found to be positive and statistically significant, suggesting that forecasts of GDP growth are 0.13 percentage points too low on average when the target year is a recovery year. Model 4 (M4) allows for richer interactions between the phases of the business cycle and the forecast errors by estimating the parameters δ_1 and δ_2. In M4 the estimated β_1 parameter is three times larger in absolute value than in M2 or M3: forecasts of a recession year during the last expansion year are biased upwards by more than 2.5 percentage points on average. When the recession starts, the forecasts of a recession year have a much smaller upward bias ($\beta_1 + \delta_1$), around half a percentage point: forecasters are initially surprised by the recession but once it hits the economy they revise their forecasts significantly lower. While the economy is in the last year of a recession, forecasts of the following year (a recovery year) are too pessimistic. The bias ($\beta_2 + \delta_1$) is around half a percentage point. Once the recovery has started, forecasts of the recovery year show a smaller bias ($\beta_2 + \delta_2$), less than three tenths of a percentage point. Again, forecasters are surprised by the recovery but, once the recovery starts, they revise their forecasts upwards.

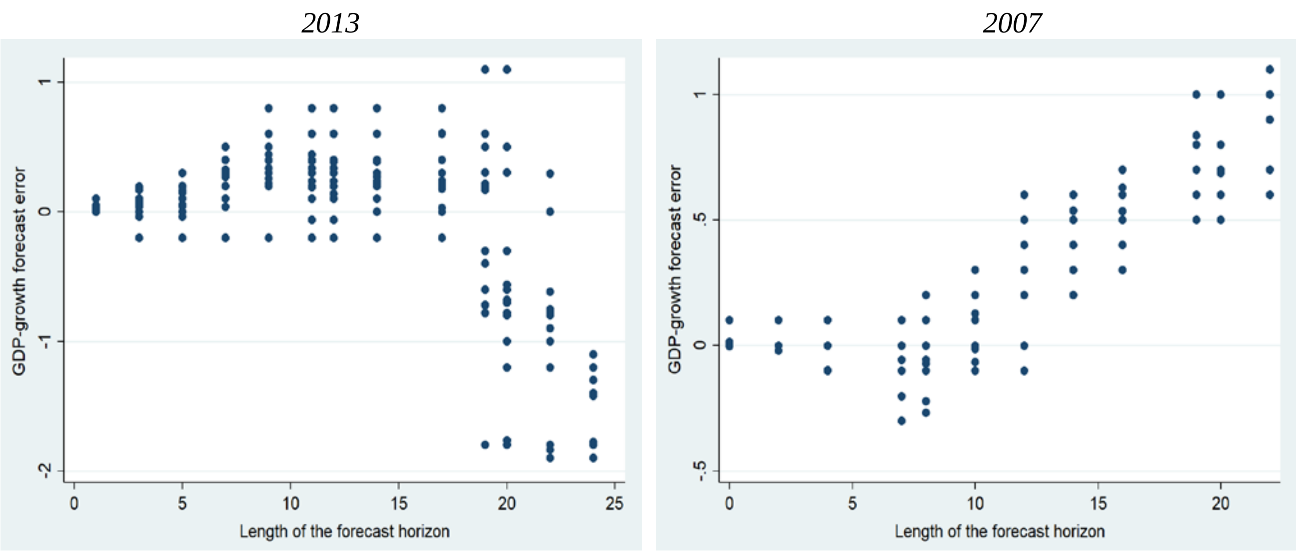
M2, M3 and M4 allow for level shifts in the forecast errors as the economy moves through the business cycle. However, Figure 2 showed that the absolute value of the average forecast error increases with the length of the forecast horizon when the business cycle is controlled for. Therefore, Model 5 (M5) estimates equation (3) without restrictions, introducing the effects of the horizon lengths in two ways: a linear effect (γ_1) and interaction effects with the dummies of the phases of the business cycle (γ_2, γ_3, γ_4 and γ_5). The linear effect is found to be statistically lower than zero but economically insignificant, one tenth of a percentage point for a horizon length of 10 months. Put differently, the forecasts have a negligible upward bias during expansions. Besides, all the effects of the business cycle on the forecast errors described in M4 are captured in M5 by the interaction effects with the length of the forecast horizon. γ_2 is negative and large (forecasters do not anticipate the recessions), γ_4 is positive but smaller in absolute value than γ_2 (forecasters revise their forecasts downwards once the recession has started but the revisions are too mild to eliminate the forecast error, as found by An, Jalles and Loungani, 2018, and Galbraith and van Norden 2019), γ_3 is negative but smaller in absolute value than γ_2 (forecasters do not anticipate the recovery), and γ_5 is positive and slightly larger than γ_4 (forecasters are slow to revise their forecasts upwards once the recovery has started).

Models M1 to M5 were estimated by Dovern and Jannsen with forecasts errors of GDP growth from the Consensus Economics Survey and the results obtained here are fully consistent with theirs. Annex IV shows the estimation results for the forecast errors of the subcomponents of GDP: private consumption, investment, government expenditure, exports and imports. For the forecast errors of the growth rate of private consumption, investment, exports and imports, the results are qualitatively similar to those obtained with GDP growth. Unsurprisingly, the results for government expenditure are somewhat different: forecasters predict growth rates of government expenditure that are too low during expansions and the bias increases with the length of the forecast horizon (γ_1 = 0.037), while this bias is negative for the rest of the variables.

3.2. The impact of past growth rates of real GDP on subsequent forecast errors

Dovern and Jannsen assumed in their model that the effect of the length of the forecast horizon on the forecast error is linear. However, this assumption fails to capture dynamics like those presented in Figure 3. The chart on the left shows all the available individual forecasts of GDP growth for target year 2013. Each dot is a different forecast. The dots closer to the right edge of the chart represent forecasts that were submitted earlier and thereby have a longer forecast-horizon length, while those closer to the left edge were submitted later and their horizon length is shorter. The relationship between the forecast error for target year 2013 and the length of the forecast horizon is clearly non-linear. The forecasts with the longest horizon lengths, which were submitted at the beginning of 2012, turned out to be too optimistic (both 2012 and 2013 were recession years). These forecasts gave rise to large negative errors. But by the end of 2012 most forecasters had changed their minds and were very pessimistic about GDP growth in 2013, so pessimistic that their forecasts were biased downwards (positive forecast errors). This evidence is at odds with the assumption of a linear relationship between forecast errors and the length of the forecast horizon.

Figure 3. Individual forecast errors for real GDP growth by horizon lengths for target years 2013 and 2007.



This hump-shaped non-linearity does not appear in recession years only. The right-hand side of the chart shows the individual forecasts of GDP growth for 2007, an expansion year. Forecasters were too pessimistic about 2007 at the beginning of 2006, resulting in positive forecast errors. But most of the forecasters became a little too optimistic by spring 2007, and their predictions exhibited negative forecast errors. This J-shaped pattern is not consistent with the assumption that the effect of the length of the forecast horizon on the forecast error is linear.

Similar examples can be found for the subcomponents of GDP. Annex IV shows hump-shaped relationships between forecast errors and the length of the forecast horizon for consumption in 2003, investment in 2013, government expenditure in 2012, exports in 2006 and imports in 2013. The model presented in the previous subsection needs to be extended to capture these patterns.

The simplest extension would consist on adding the length of the forecast horizon squared as an additional regressor, yielding a quadratic relationship between the forecast error and the horizon length. I will come back to that extension shortly but first let me investigate a more interesting alternative: could it be the case that professional forecasters are backward-looking and put too much weight on the most recent macroeconomic developments? On the left-hand side of Figure 3 I discussed that forecasters in early 2012 were too optimistic about GDP growth in 2013. The reason could be that the first estimate of GDP growth in 2011, published in February 2012, came in at 0.7%. As 2010 was believed to be a recession year at the time, it seemed that the recession triggered by the financial crisis was finally over. The positive news might have led to excessive upward revisions to GDP-growth forecasts. The large negative forecast errors for horizon lengths between 22 and 24 months are consistent with this hypothesis.

The optimism, however, was short-lived. In September 2012, the growth rate of GDP in 2011 was revised down to 0.4%, and it was revised down again to 0.0% in September 2013. The final figure came in as low as -1.0%. Moreover, the first estimate of GDP growth in 2012, published in March 2013, was dismal: -1,4%. The hopes of an end to the recession vanished with the eruption of the sovereign debt crisis in the euro area. In this context, most professional forecasters might have become too pessimistic about growth in 2013, they revised their forecasts downwards too much and, as a result, they made positive forecast errors for horizon lengths lower than 17 months.

To test for this kind of irrational backward-lookingness or excessive influence of the latest macroeconomic data releases on the forecasts and thereby on the forecast errors, I first investigate the relationship between past shocks and subsequent forecast errors with a simple theoretical example. Let's assume that there are three periods in the model: period 0, when the economy is in its steady state, period 1, when a shock hits the economy, and period 2, the target period of the forecasts. In other words, agents try to predict the state of the economy in period 2 after the period 1 shock. Let's assume that the target variable is GDP growth, in deviations from steady state, which follows an AR(1) process:

$$\Delta GDP_t = \rho \Delta GDP_{t-1} + \varepsilon_t \qquad (4)$$

If agents were irrationally backward-looking and used a random walk to forecast the growth rate of GDP in period 2, instead of (4),

$$E_1 \Delta GDP_2 = \Delta GDP_1 \qquad (5)$$

the forecast error would then be:

$$\Delta GDP_2 - E_1 \Delta GDP_2 = (\rho - 1) \Delta GDP_1 + \varepsilon_2 \qquad (6)$$

Therefore, the regression coefficient of forecast errors on past GDP growth would be negative as long as GDP growth were stationary.

If agents were rational but inattentive, they would observe the shock ε_1 with an individual probability τ ($0 < \tau < 1$). If they observed the shock they would update their expectations optimally:

$$E_1 \Delta GDP_2 = \rho \Delta GDP_1 \qquad (7)$$

And their forecast errors would be unpredictable. But with probability 1- τ they would not observe the shock and they would not update,

$$E_1 \Delta GDP_2 = 0 \qquad (8)$$

because the GDP growth was in its steady state in period 0. For the agents that do not update, their forecast error in period 2 is:

$$\Delta GDP_2 - E_1\Delta GDP_2 = \rho\Delta GDP_1 + \varepsilon_2 \quad (9)$$

In this case, if agents were rationally inattentive, the regression coefficient of forecast errors on past GDP growth would be positive as long as GDP growth were persistent.

Finally, let's consider the case when agents are rational but receive a noisy signal in period 1 (assume for simplicity that they observe GDP growth at period 0):

$$\Delta GDP_{1i} = \Delta GDP_1 + \eta_{1i} \quad (10)$$

where η_{1i} is a zero-mean disturbance, uncorrelated to ΔGDP_1 and interpreted as the noise that receives agent *i*, The forecast of GDP growth in period 2 would be as follows:

$$E_{1i}\Delta GDP_2 = (1-\alpha)E_0\Delta GDP_2 + \alpha\rho(\Delta GDP_1 + \eta_{1i}) \quad (11)$$

where α is the Kalman gain. If there were no noise, $1-\alpha = \eta_{1i} = 0$, the growth rate of GDP in period 1 would be observable and the expected forecast error would be zero (the first term on the right side of equation (11) is zero because there were neither shocks nor noise before period 1). As long as there were noise in the signal the forecasters received in period 1, their forecast errors would be:

$$\Delta GDP_2 - E_{1i}\Delta GDP_2 = (1-\alpha)\rho\Delta GDP_1 - \alpha\rho\eta_{1i} + \varepsilon_2 \quad (12)$$

If agents were rational but received noisy information, the regression coefficient of forecast errors on past GDP growth would be positive as long as GDP growth is persistent ($\rho > 0$).

Following the intuition from this simple theoretical example I explore the empirical relationship between forecast errors and past growth rates of GDP. To this end, I start from Dovern and Jannsen's model as a baseline because the forecast horizon length of the predictions in the Funcas database is not always one and the horizon length could affect the size of potential biases (Ager *et al*., 2009). Then, I add eight additional regressors: the real-time estimates of quarterly GDP growth rates over the previous eight quarters for which data was available at the time of producing each forecast.[10] These regressors are predetermined at the time of producing the forecasts:

$$FE_{i,GDP,t,h} = \Sigma_i\,\alpha_i D_i + (\beta_1 + \gamma_2\,h)\,D_{recession} + (\beta_2 + \gamma_3\,h)\,D_{recovery} + (\delta_1 + \gamma_4\,h)\,D_{recession\ started} +$$
$$+ (\delta_2 + \gamma_5\,h)\,D_{recovery\ started} + \gamma_1\,h + \Sigma_j\,\lambda_j\Delta GDP_j + \varepsilon_t \qquad for\ i=1,\ldots,23\ and\ j=1,..,8 \quad (13)$$

ΔGDP_1 is the real-time estimate of quarterly GDP growth for the most recent quarter available at the time of producing the forecast. ΔGDP_2 is the real-time estimate of quarterly GDP growth for the quarter before the most recent, and so on up to ΔGDP_8. If forecasters are irrationally backward-looking, the estimated λ_j should be statistically lower than zero.[11] If forecasters are rational but inattentive, the estimated λ_j should be statistically larger than zero. If forecasters are rational but receive noisy information about GDP growth, the estimated λ_j should be statistically larger than zero. If forecasters are rational and there are no information rigidities, the estimated λ_j should not be statistically different from zero.

The estimation results are displayed in column 6 of Table 1 (model M6). The coefficients of six out of eight past GDP growth rates are negative, four of them statistically lower than zero. The sum of the eight λ_j is negative and statistically significant. And the effects are economically relevant: a quarterly GDP growth rate of 1% above

[10] The Akaike and Bayesian information criteria selected 8 lags as the more informative specification conditional on a maximum of 8 lags to avoid losing too many observations.

[11] Capistran and Timmerman (2009) built a model where forecasters have asymmetric loss functions with heterogeneity. In this model, Coibion and Gorodnichenko (2015) showed that past data may be negatively correlated with subsequent forecast errors if the degree of asymmetry is non-zero on average. But if that were the case average forecast errors during expansions would not remain as close to zero as shown in Figure 2.

its long term value causes a forecast error of -0.28 percentage points one quarter later, -0.49 percentage points one year later and -1.04 percentage points two years later.

It could be argued that this result may be driven by the presence of several recessions in a relatively short sample: *conditional on big negative shocks to GDP growth,* the correlation between past GDP growth rates and the following forecast errors could be negative not only for backward-looking forecasters but also for rational forecasters (with or without information rigidities).[12] To check whether the negative effect of GDP-growth rates on subsequent forecast errors found in model M6 also hold during expansions, I run the same regression with interaction terms between past GDP growth rates and the dummy variables $D_{expansion}$, which is equal to 1 if the target year is an expansion year and 0 otherwise, $D_{recession}$ and $D_{recovery}$:

$$FE_{i,GDP,t,h} = \Sigma_i\, \alpha_i D_i + (\beta_1 + \gamma_2\, h)\, D_{recession} + (\beta_2 + \gamma_3\, h)\, D_{recovery} + (\delta_1 + \gamma_4\, h)\, D_{recession\ started} + $$
$$+ (\delta_2 + \gamma_5\, h)\, D_{recovery\ started} + \gamma_1\, h + \Sigma_j\, \lambda_{j,exp} D_{expansion}\, GDP_j + \Sigma_j\, \lambda_{j,rece} D_{recession}\, GDP_j +$$
$$+ \Sigma_j\, \lambda_{j,reco} D_{recovery}\, GDP_j + \varepsilon_t \qquad for\ i=1,\dots,23\ and\ j=1,..,8 \qquad (14)$$

There are 24 λ_j to be estimated in this model. Therefore, Table 1 only reports its sum for each phase of the business cycle (model M7 in Table 1). These sums are statistically lower than zero not only for recessions but also for expansions.[13] The results are consistent with the hypothesis that the forecasters are backward looking and overreact irrationally to short-term macroeconomic developments: when GDP growth is high they revise up their forecasts too much and they make negative forecast errors; and when GDP growth is relatively low, they revise down their forecasts too much and they make positive forecast errors. These overreactions occur not only in recessions, when the estimated effects are stronger, but also for expansions.

Another robustness check is done by adding to the right-hand side of equation (13) the quadratic horizon effects I mentioned above:

$$FE_{i,GDP,t,h} = \Sigma_i\, \alpha_i D_i + (\beta_1 + \gamma_2\, h + \gamma_{2squared}\, h^2)\, D_{recession} + (\beta_2 + \gamma_3\, h + \gamma_{3squared}\, h^2)\, D_{recovery} +$$
$$+ (\delta_1 + \gamma_4\, h + \gamma_{4squared}\, h^2)\, D_{recession\ started} + (\delta_2 + \gamma_5\, h + \gamma_{5squared}\, h^2)\, D_{recovery\ started} +$$
$$+ \gamma_1\, h + \gamma_{1squared}\, h^2 + \Sigma_j\, \lambda_j\, GDP_j + \varepsilon_t \qquad for\ i=1,\dots,23\ and\ j=1,..,8 \qquad (15)$$

The results are shown in column 8 of Table 1 (model M8) and are qualitatively and quantitatively very similar to those of M6, with most of the estimated λ_j statistically lower than zero. Thus, in all three models, M6, M7 and M8, there is evidence that forecasters may have behaved in a backward-looking manner by revising their forecasts too aggressively in response to short-term macroeconomic developments. The same evidence of backward-lookingness and overreaction to recent data releases is found for all subcomponents of GDP, a result that suggests this feature does not depend on the series to be predicted (see Annex V).

The results are in line with recent findings reported by Gelain *et al.* (2019), who found that SPF forecasts are alike to forecasts from a DSGE model with irrational agents that place too much weight on the most recent observations when computing their expectations. They also support the findings by Andrade and Le Bihan (2013), who pointed out that a model with rational inattention and noisy information cannot generate the large degree of persistence of forecast errors observed in the ECB's SPF. The results are consistent with those reported by Ang *et al*. (2007), who found that inflation forecasts from the Livingston survey are too high after high-inflation episodes and too low after low-inflation ones. The results are also closely related to the "excessive optimism" displayed by stock-market investors when they forecast stock prices at market peaks (Adam *et al.*, 2017, Greenwood and Schleifer, 2014). Finally, the results are in line with the "available heuristic" in

[12] See, for example, Coibion and Gorodnichenko (2015): "rejections of the null of full-information rational expectations are much more common over short samples in which specific episodes, such as the Volcker disinflation, can have a disproportionate influence on measuring the predictability of forecast errors.".

[13] For recoveries, the sum of the λ_j is positive but, for the reason explained at the end of the previous section, the uncertainty around their point estimates is relatively large: only three of them are statistically different from zero, adding up to -0.125).

psychology, according to which agents put too much weight on recent events because they are "easily accessible" (Koursaros, 2018).

Coibion and Gorodnichenko (2012) found that aggregate forecast errors of inflation from the US SPF are not correlated to past inflation levels, a result that is at odds with the results presented here. I believe there are several differences between their paper and this paper that may help explain this discrepancy. First, different surveys are used. Second, different variables are predicted. Third, they used average forecasts while individual forecasts are considered here. And fourth, they only analysed one forecast horizon length (12 months) while many different horizons are used here. This last difference is, in my view, the most relevant because Figure 2 shows that forecast errors are much larger in absolute terms for forecast horizon lengths longer than 12 months, and Ager *et al.* (2009) found that professional forecasts are more likely to be biased at horizon lengths longer than 12 months. Interestingly, Coibion and Gorodnichenko (2012) also estimated the effect of past levels of inflation on forecasts errors of inflation using the Livingston survey and a survey of forecasts by members of the Federal Open-Market Committee. The estimated impact in both cases is negative and statistically significant, as described in the present paper. They also obtained negative estimates of the coefficients of past inflation using US SPF data when semi-annual forecasts were used instead of quarterly forecasts.

Hubert and Mirza (2019) used the New Keynesian Phillips Curve framework to find that inflation nowcasts from the US SPF can be characterised as a weighted average of past inflation and inflation forecasts one quarter ahead. They concluded that professional forecasters are forward-looking because the coefficient of past inflation is small (0.2). However, as they also indicated, if forecasts of inflation one quarter ahead embedded information from the past, professional forecasters would be more backward-looking than their estimates suggest. As a matter of fact, when they replaced one-quarter-ahead forecasts with four-quarter-ahead forecasts, the weight of the backward-looking component increased to 0.6.

3.3 Goodness of fit of the extended model vs. the original model

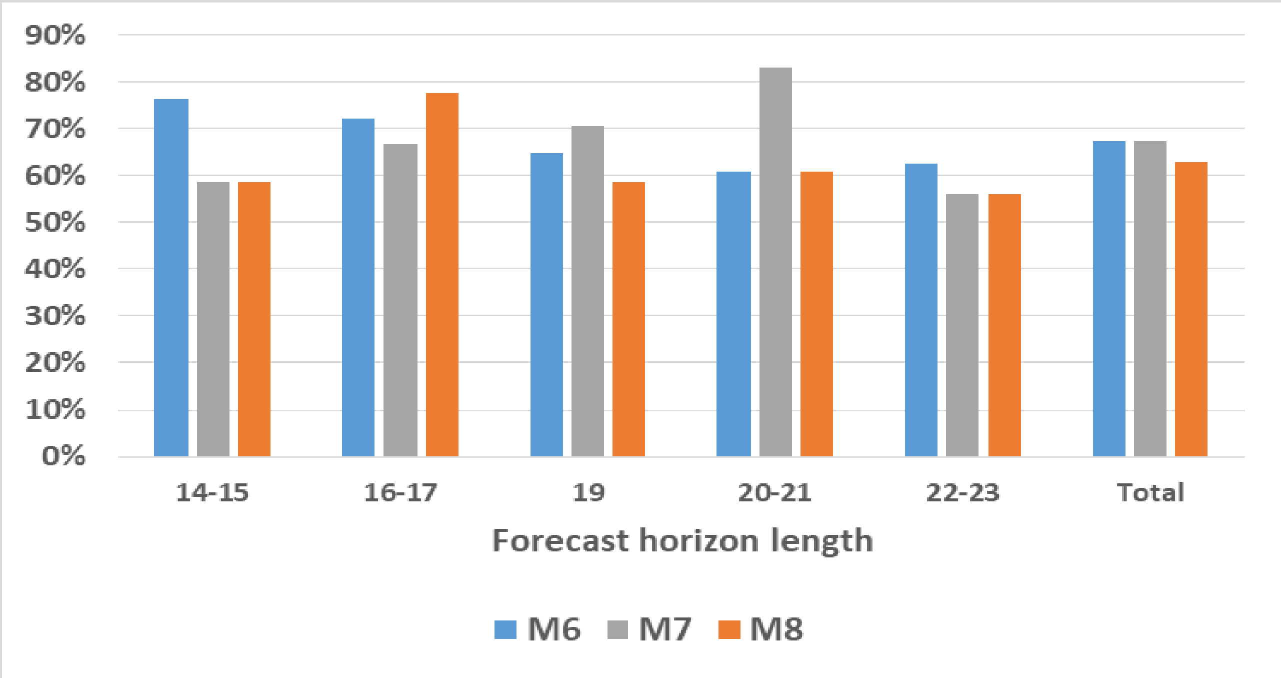
The previous subsection has shown that recent macroeconomic developments help explain forecast errors better in a statistical sense. But how big is this improvement in forecasting terms? To answer this question I conducted a horse race between model M5 (the estimated original model) on the one hand and the three extensions estimated in this paper, M6, M7 and M8, on the other.[14] The horse race consists on counting how many times the fitted forecast errors from M6, M7 and M8 are closer to the consensus forecast errors than the fitted forecast errors of the estimated original model. To make the exercise more interesting I focus on forecast horizon lengths longer than 13 months as, not surprisingly, these are the lengths for which the largest forecast errors are made.

The results are displayed in Figure 5. The extended models beat the original model more than 50% of the time for every single horizon longer than 13 months. Putting all these horizon lengths together, M6 and M7 outperform the original model M5 in 67.4% of the comparisons while M8 does it in 62.7%. Very similar results are obtained when the horse race is conducted for the forecast errors of the subcomponents of GDP (see Annex VI). The best performing model is the most parsimonious extended model, M6. Its success rate versus the original model is 77% for consumption, 65% for investment, 57% for government expenditure, 83% for exports and 72% for imports.

These results have important implications for the Funcas survey. They imply that the average or consensus forecasts published by Funcas could be improved upon by predicting the forecast errors with models M5 to M8. While all the extended models presented here are useful for this task, M6 is the model that does this job the best most of the time because it captures how professional forecasters seem to overreact to recent releases of macroeconomic data: when recent data releases are relatively good, *ceteris paribus*, forecast errors become more negative, and when recent data releases are relatively bad, forecast errors become more positive. These systematic backward-lookingness may be exploited by Funcas to improve the quality of its consensus forecasts.

[14] A more formal Diebold-Mariano test is not conducted for two reasons. From a theoretical perspective, the Dieblod-Mariano test should not be used to compare nested models because the numerator and denominator of the test statistic asymptotically converge to zero under the null hypothesis. From an empirical perspective, there are too few observations available for each forecast-horizon length.

Figure 5. Percentage of times each extended model outperforms the original model when predicting the consensus forecast errors of GDP growth.



4. The empirical model of the forecast revisions

The results presented in the previous section are consistent with backward-lookingness or overreaction to the recent past by professional forecasts. As described above, the mechanism could work as follows: when the latest releases of macroeconomic data are relatively good, forecasters may become too optimistic, revise their forecast upwards too much and make negative forecast errors. While the link between data releases and forecast errors has been documented in Section 3, I have not shown any evidence of the link between data releases and forecast revisions yet.

More importantly, there is a feature of the forecast errors of real GDP growth that cannot be explained by backward-lookingness alone: their persistence. Forecast errors typically display positive autocorrelation (Capistran and Timmermann, 2009, Bowles *et al.* (2010), Andrade and Le Bihan, 2013, An *et al.*, 2018, Galbraith and van Norden, 2019). If professional forecasters overreacted to high growth rates at the end of an expansion, became too optimistic and made negative forecast errors, there is no reason to remain too optimistic once the recession hits the economy. Consequently, the time series of forecast errors could quickly converge to zero when forecasters are backward-looking, or could even display negative autocorrelation as forecasters could produce forecasts that are too pessimistic in response to the onset of the recession.

The theories of rational inattention and rational forecasting with noisy signals were developed in part to address this issue. Inattentive forecasters may generate persistence in aggregate forecast errors because not all forecasters update their forecasts at the same time. Rational forecasters that receive noisy signals may react too timidly on average to the signal and, as more information is revealed, they slowly revise their forecasts in the same direction during multiple periods. Empirically, however, models incorporating these features are not able to generate enough persistence in forecast errors (Andrade and Le Bihan, 2013).

Therefore, there must be another reason for the inertia of forecast errors over time. To explore this issue, the next step is to estimate a model of forecast revisions. The estimated model is:

$$FREV_{i,GDP,t\text{-}k,t} = \alpha + \Sigma_i\, \alpha_i D_i + \rho\, FREV_{i,GDP,t\text{-}k\text{-}x,t\text{-}k} + (\beta_1 + \beta_2 D_{2010:9}) LFE_{i,qGDP,t} + \Sigma_j\, \lambda_j REV_{GDP,\text{-}j,t\text{-}k\text{-}1,t\text{-}1} +$$
$$+ \gamma_1(F_{i,GDP,t\text{-}k} - F_{consensus,GDP,t\text{-}1}) + \gamma_2(F_{i,GDP,t\text{-}k} - F_{government,GDP,t\text{-}1}) +$$
$$+ \gamma_3(F_{i,GDP,t\text{-}k} - F_{EUCommission,GDP,t\text{-}1}) + \gamma_4(F_{i,GDP,t\text{-}k} - F_{IMF,GDP,t\text{-}1}) + \gamma_5(F_{i,GDP,t\text{-}k} - F_{OECD,GDP,t\text{-}1}) + \varepsilon_t$$
$$\text{for } i=1,\ldots,23 \text{ and } j=1,..,4 \qquad (16)$$

$FREV_{i,GDP,t\text{-}k,t}$ is the revision at time t of a forecast last submitted in t-k. The forecast is submitted by forecaster i and refers to GDP growth for a calendar year; α is a constant. The α_i are individual effects. $FREV_{i,GDP,t\text{-}k\text{-}x,t\text{-}k}$ is the latest revision by forecaster i before time t. It is included in the model because An, Jalles, and Loungani, (2018) found that revisions to GDP-growth forecasts display inertia and thereby are autocorrelated. $LFE_{i,qGDP,t}$ is the latest GDP-growth forecast error by forecaster i at the time of the revision of the forecast. I expect β_1 to be positive because better than expected realisations of GDP growth may trigger upward revisions to GDP forecasts. This expectation is based on the relationship between forecast errors of quarterly GDP growth and subsequent revisions to GDP-growth forecasts in the Funcas survey shown in Figure 6.

The revisions used in this section are the changes to GDP-growth forecasts for calendar years, the same forecasts used in the previous section, from one survey round to the next, in percentage points. This notwithstanding, in this section I am not using the forecast errors employed in Section 3 because they were computed from forecasts of GDP growth for calendar years and, if I used them, I would only have one realisation per year. Instead, I compute forecast errors of *quarterly* GDP growth because Funcas also collects individual forecasts for this variable in its survey six times a year, although these forecasts are not published. These forecasts refer to year-on-year quarterly growth rates before September 2010 and to quarter-on-quarter growth rates thereafter. Unfortunately, Funcas was not able to provide these forecasts after July 2013 because it has not stored the quarterly forecasts since then. Therefore, the sample used in this section goes from May 1999 to July 2013.[15]

Figure 6: Relationship between forecast errors of quarterly GDP growth and revisions to GDP forecasts.

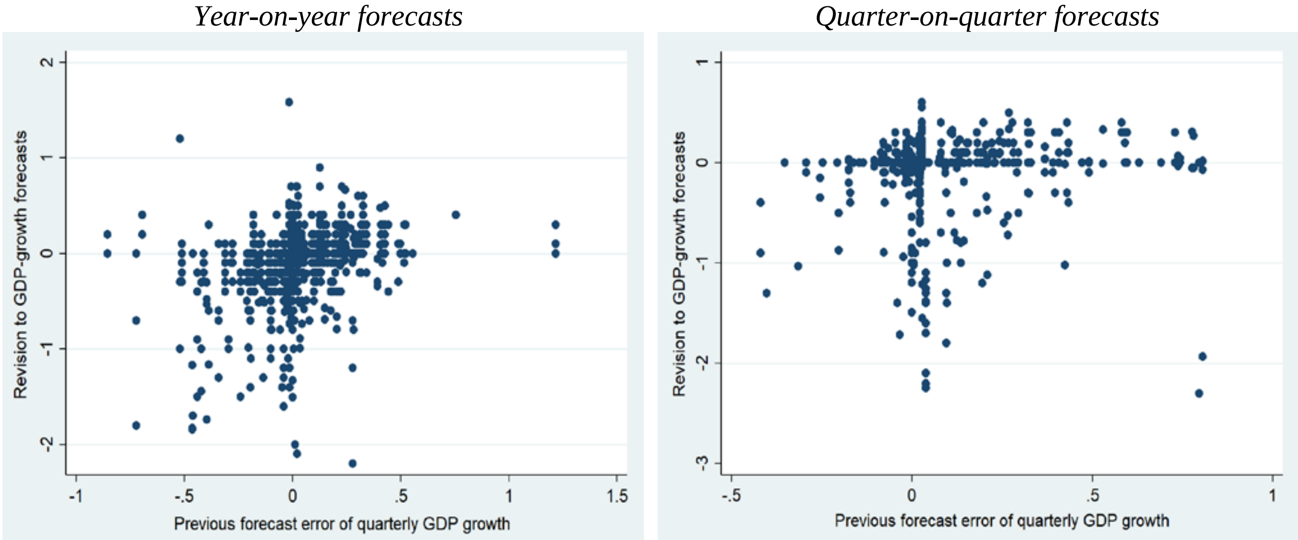


Figure 6 shows that there may be a positive relationship between forecast errors and subsequent forecast revisions, especially for the year-on-year forecasts: when quarterly forecasts are higher than the realisations, the forecasts for the calendar years are revised downwards. This relationship seems to be a bit clearer for forecasts of the growth rate of consumption (see Annex VII). There are also tentative indications of inattentive forecasters, especially for quarter-on-quarter forecasts of GDP growth, because revisions are frequently zero even after relatively large forecast errors.

[15] The quarterly forecasts submitted for the September 2008 and November 2009 surveys were not available either.

To capture the shift in September 2010 from year-on-year forecasts of quarterly GDP growth to quarter-on-quarter forecasts, a dummy variable, $D_{2010:9}$, is included in equation (16). It is 0 before September 2010 and 1 thereafter. I expect β_2 to be negative and offset β_1 at least partially, if not totally, because forecast errors of quarter-on-quarter growth rates may be attributed more to transitory short-term volatility than to persistent shifts in medium-term growth, and thus may trigger smaller revisions, if any, to forecasts of calendar years (Croushore, 2010).

$REV_{GDP,-j,t-k-1,t-1}$, with j from 1 to 4, are the revisions between $t-k$ and $t-1$ to the latest four realisations of the level of quarterly GDP, in percentage points, before each forecast revision. I expect that the four λ_j may be positive, because positive revisions to past GDP may be an indication that the economy is in better shape than previously assumed and may trigger upward revisions to forecasts of future GDP growth. Revisions to released GDP data are obtained from the OECD real-time database. To avoid mixing revisions to released GDP data, which may trigger forecast revisions by forecasters, with methodological changes that may cause jumps in revised data, which are unlikely to cause forecast revisions, I define methodological changes as revisions of more than 10 percentage points to GDP data. Revisions of more than 10 percentage points to released GDP data took place in February 2002 (the euro changeover) but also in September 1999, September 2002, June 2005 and December 2011. These very large revisions are excluded from the analysis.[16]

$F_{i,GDP,t-k} - F_{consensus,GDP,t-1}$ is the difference between the previous GDP-growth forecast by forecaster i and the latest consensus forecast available, i.e. the average forecast from the previous survey round. If, for strategic reasons, forecasters try to submit predictions that are relatively close to the consensus forecast, a behaviour commonly known as *herding*, γ_1 should be negative (Morris and Shin, 2002, Coibion and Gorodnichenko, 2012). Rational forecasters have no reason to herd. Therefore, if the forecasters in the Funcas database were rational but inattentive or were rational forecasters with noisy information, γ_1 would be zero. But if the forecasters displayed some form of irrationality, e.g. fears of making large forecast errors when everybody else is forecasting accurately, herding could appear. Importantly, herding around the consensus forecast may amplify the inertia of both individual and average forecast errors (Coibion and Gorodnichenko, 2012), a feature that could explain, if $\gamma_1 < 0$, why backward-looking forecasters that overreact to recent data releases make persistent forecast errors.

The forecasters surveyed by Funcas could also display herding behaviour around the latest forecast of GDP growth available from the Spanish government ($F_{government,GDP,t-1}$), from the European Commission ($F_{EUCommission,GDP,t-1}$), from the International Monetary Fund ($F_{IMF,GDP,t-1}$) or from the OECD ($F_{OECD,GDP,t-1}$).[17] If that were the case, γ_2, γ_3, γ_4, γ_5 would be lower than zero.

The information assumption made here is that all variables published before the month of a Funcas survey round are known to the forecasters participating in the round. All the information available during the month of the survey round is assumed to be unknown to the survey participants until the next survey round.

The results of estimating equation (16) by pooled OLS with bootstrap standard errors are shown in the first three columns of Table 2.[18] The results in the first column are the estimated coefficients when the individual effects, α_i, are zero and there is no inertia in forecast revisions, ρ=0 . In the second column, the model includes individual effects and the constant, α, is excluded, but there is still no inertia in forecast revisions. In the third column, the model allows for individual effects and inertia in forecast revisions. In all three cases, as expected, the estimated β_1 is positive, statistically significant and economically relevant: when professional forecasters receive a one-percentage-point surprise to the growth rate of year-on-year quarterly GDP they revise up their forecasts for the calendar year by around 0.4 percentage points. Furthermore, $\beta_1 + \beta_2$ is statistically equal to zero, which suggest

[16] Methodological changes to the private consumption series are identified to occur in September 1999, February 2002, April 2003, June 2005 and December 2011.
[17] All the forecasts from the Spanish government and international organisations were also provided by Funcas.
[18] Bootstrap standard errors are needed because model errors are heteroskedastic (more variance in more volatile years) and cross-correlated (revisions of forecasts one and two calendar years ahead have typically the same sign).

that the revisions by professional forecasters are not triggered by a surprise in just one quarter but by a sequence of news over several quarters.[19]

Table 2. Estimation results of the empirical model of forecast revisions.

	GDP (No individual effects)	GDP (Individual effects)	GDP (Individual effects)	C (No individual effects)	C (Individual effects)	C (Individual effects)
ρ	-	-	0.295	-	-	0.001
			(0.00)			(0.98)
β_1	0.427	0.441	0.334	0.200	0.221	0.277
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
β_2	-0.479	-0.488	-0.484	-0.303	-0.338	-0.344
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
λ_1	0.080	0.067	-0.017	-0.107	-0.119	-0.143
	(0.28)	(0.36)	(0.81)	(0.40)	(0.36)	(0.26)
λ_2	0.087	0.084	0.067	-0.089	-0.087	-0.098
	(0.15)	(0.16)	(0.25)	(0.45)	(0.47)	(0.40)
λ_3	-0.254	-0.258	-0.221	0.289	0.281	0.313
	(0.00)	(0.00)	(0.00)	(0.05)	(0.06)	(0.04)
λ_4	-0.077	-0.063	-0.022	0.156	0.178	0.196
	(0.35)	(0.45)	(0.78)	(0.11)	(0.08)	(0.04)
γ_1	-0.684	-0.711	-0.774	-0.657	-0.689	-0.687
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
γ_2	-0.092	-0.090	-0.036	0.088	0.091	0.097
	(0.00)	(0.00)	(0.23)	(0.00)	(0.00)	(0.00)
γ_3	-0.148	-0.152	-0.144	-0.071	-0.075	-0.059
	(0.00)	(0.01)	(0.01)	(0.07)	(0.05)	(0.14)
γ_4	0.003	0.006	-0.003	-0.267	-0.266	-0.265
	(0.95)	(0.90)	(0.95)	(0.00)	(0.00)	(0.00)
γ_5	-0.046	-0.043	-0.016	-0.051	-0.045	-0.042
	(0.31)	(0.35)	(0.72)	(0.08)	(0.11)	(0.17)
α	-0.031	-	-	-0.037	-	-
	(0.00)			(0.00)		
Observations	1728	1728	1607	1289	1289	1188
Adj. R^2	0.21	0.24	0.34	0.31	0.35	0.36

Note: P-values in parenthesis.

The second interesting result is that there is strong evidence of herding behaviour around the consensus forecast, as γ_1 is negative, statistically significant and very large. When the consensus forecast is above an individual forecast by one percentage point, the forecaster who submitted such forecast revises it up by around 0.7 percentage points in the next survey round.[20] This feature helps explain persistent forecast errors in the presence of forecasters that overreact to the recent past. There seem to be herding effects around the forecasts submitted by the European Commission as well, but the quantitative effects are negligible.

[19] The p-values of the F-test of the null hypothesis $\beta_1 + \beta_2 = 0$ are 0.60, 0.59 and 0.10 for the models in columns 1, 2 and 3 respectively.

[20] Strong herding behaviour is also found by Döpke, Fritsche and Waldhof (2019) using a German survey of professional forecasters.

Another finding not consistent with forecast rationality is the statistical significance of previous forecast revisions in column 3, Revisions of rational forecasters should be uncorrelated with the information available at the time of the previous revision, but these results show that positive revisions are more likely to be followed by other positive revisions rather than by negative revisions. Finally, revisions to previously-published GDP data do not seem to have a major impact on forecast revisions.

The results for the model of revisions to forecasts of the growth rate of private consumption are displayed in columns four to six in table 2. All the conclusions drawn for revisions of GDP forecasts apply here with three exceptions. First, the size of the estimated β_1 coefficient is a bit smaller than in the model of revisions of GDP forecasts, which makes sense because consumption is typically more persistent than GDP. Second, herding around the IMF forecasts of consumption growth is statistically significant and sizable. And third, there is no inertia in the revisions to forecasts of the growth rate of consumption.

All in all, the findings obtained from the model of forecast revisions, in particular the strong herding behaviour around the consensus forecast, are consistent with irrational behaviour by some professional forecasters. Putting these findings together with those from the previous section, it could be argued that the inability by professional forecasters to predict recessions could be partly related to (i) the excessive backward-lookingness and overreaction to past data, and (ii) a strong herding behaviour. Excessive backward-lookingness may lead to big negative forecast errors at the beginning of recessions. The strong herding behaviour may prevent the forecasters perceiving the early signs of a recession from deviating too much from the consensus forecast. These results help explain the persistence of forecast errors, and are consistent with the findings by Galbraith and van Norden (2019) and by An, Jalles and Lougani (2018), who observed that forecasters adjust their expectations when a recession is coming but not as much as needed to predict it accurately.

5. Conclusions

This paper investigates why professional forecasters are not effective in predicting economic recessions. To that end, I have used a Spanish survey of professional forecasters which allows to test whether the well-documented lack of success of the consensus forecast in predicting recessions is found using individual data as well. I also analyse if professional forecasters are more successful predicting the subcomponents of GDP: private consumption, investment, government expenditure, exports and imports.

The main findings of the paper may be summarised as follows. Individual forecasters are indeed too optimistic before recessions. In other words, the results reported by Dovern and Jannsen are also found with individual data. This finding applies to almost all subcomponents of GDP, no matter how volatile they are, suggesting that something more fundamental is at the root of the inability to predict the turning points of the cycle. Two factors may be playing a role in this regard. First, forecasters may be putting too much weight on past data (irrational backward-lookingess), which may lead to excessive optimism just before a recession occurs. Second, there seems to be evidence of a strong herding behaviour around the consensus forecast, which may prevent those forecasters perceiving the early signs of a recession from adjusting their expectations as much as needed to predict it.

These findings have important implications for professional forecasters and for the institutions that survey them. The hypothesis that agents are rational but are subject to information rigidities, either rational inattention or noisy information, does not find support in the data. Instead, professional forecasters could improve their forecasting performance, first, by being more forward-looking and avoiding overreactions to recent data releases, and second, by not herding around the consensus forecast. The results of this paper suggest that these improvements will increase the chances of predicting recessions more accurately. Institutions like Funcas that run surveys of professional forecasters may want to survey density forecasts or probability distributions in addition to point forecasts. Density forecasts allow forecasters to report the probabilities of tail events, like the start of a recession. They might also remove the focus from the point forecasts, which are commonly perceived as a prediction conditional on remaining in the current phase of the cycle.

Future directions for research include extending the analysis to forecasts for other countries with an aim to check whether the findings reported here are distinctive of the Spanish economy or can also be found in other jurisdictions. I also plan to combine the multiple univariate models presented in this paper into a multivariate analysis. In a multivariate framework, for example, revisions in the forecast of a variable may depend on forecast errors from other variables. Or forecast errors in one variable may depend on recent releases of other macroeconomic variables. In this context, a multivariate analysis could help increase the goodness of fit of the model of forecast revisions, whose adjusted R^2 is relatively low.

References

Adam, K., Marcet, A., and Beutel, J. (2017). Stock price booms and expected capital gains. *American Economic Review* 107, 2352-2408.

Ager, P., Kappler, M., and Osterloh, S. (2009). The accuracy and efficiency of the Consensus Forecasts: A further application and extension of the pooled approach. *International Journal of Forecasting* 25, 167-181.

An, Z., Jalles, J.T., and Loungani, P. (2018). How well do economists forecast recessions? IMF Working Paper WP/18/39.

Andrade, P., and Le Bihan, H. (2013). Inattentive professional forecasters. Journal of Monetary Economics, 60, 967-982.

Ang, A., Bekaert, G., and Wei, M. (2007). Do macro variables, asset markets, or surveys forecast inflation better? *Journal of Monetary Economics* 54, 1163–1212.

Baghestani, H. (2019). Long-term interest rate predictability: Exploring the usefulness of survey forecasts of growth and inflation. *Cogent Economics & Finance* 7, 1582317.

Best, G., and Kapinos, P. (2018). Is the Fed's news perception different from the private sector's? *Applied Economics*, forthcoming.

Bonham, C. S., and Cohen, R. H. (2001). To aggregate, pool, or neither: testing the rational-expectations hypothesis using survey data. *Journal of Business & Economic Statistics* 19, 278–291.

Bowles, C., Friz, R., Genre, V., Kenny, G., Meyler, A., and Rautanen, T. (2010). An evaluation of the growth and unemployment forecasts in the ECB survey of professional forecasters. *Journal of Business Cycle Measurement and Analysis* 4, 1-28.

Capistran, C., and Timmermann, A. (2009). Disagreement and biases in inflation expectations. *Journal of Money, Credit and Banking* 41, 365–396.

Coibion, O., and Gorodnichenko, Y. (2012). What can survey forecast tell us about information rigidities? *Journal of Political Economy* 120, 116-159.

Coibion, O., and Gorodnichenko, Y. (2015). Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts. *American Economic Review* 105, 2644-2678.

Conflitti, C., De Mol, C., and Giannone, D. (2015). Optimal combination of survey forecasts. *International Journal of Forecasting* 31, 1096–1103.

Croushore, D. (2010). An evaluation of inflation forecasts from surveys using real-time data. *BE Journal of Macroeconomics* 10, 1–32.

Davig, T., and Hall, A.S. (2019). Recession forecasting using Bayesian classification. *International Journal of Forecasting* 35, pp. 848-867.

Deschamps, B., and Ioannidis, C. (2013). Can rational stubbornness explain forecast biases? *Journal of Economic Behavior and Organization* 92, 141–151.

Diebold, F.X., and Shin, M. (2018). Machine learning for regularized survey forecast combination: Partially-egalitarian LASSO and its derivatives. *International Journal of Forecasting*, forthcoming.

Döpke, J., Fritsche, U., and Waldhof, G. (2019). Theories, techniques and the formation of German business cycle forecasts: evidence from a survey of professional forecasters. *Journal of Economics and Statistics* 239, 203-241.

Dovern, J., Fritsche, U., Loungani, P., and Tamirisa, N. (2015). Information rigidities: comparing average and individual forecasts for a large international panel. *International Journal of Forecasting* 31, 144-154.

Dovern, J., and Jannsen, N. (2017). Systematic errors in growth expectations over the business cycle. *International Journal of Forecasting* 33, 760-769.

Dovern, J., and Weisser, J. (2011). Accuracy, unbiasedness and efficiency of professional macroeconomic forecasts: An empirical comparison for the G7. *International Journal of Forecasting* 27, 452–465.

ECB (2019). Results of the third special questionnaire for participants in the ECB Survey of Professional Forecasters. Retrieved from *https://www.ecb.europa.eu/stats/ecb_surveys/survey_of_professional_forecasters/html/ecb.spf201902_specialsurvey~7275f9e7e6.en.html#toc1* on 26 April 2019.

El-Shagi, M. (2018), Rationality tests in the presence of instabilities in finite samples. *Economic Modelling*, forthcoming.

Frenkel, M., Lis, E.M., and Rülke, J.C. (2011). Has the economic crisis of 2007–2009 changed the expectation formation process in the euro area? *Economic Modelling* 28, 1808–1814.
Galbraith, J.W., and van Norden, S. (2019). Asymmetry in unemployment rate forecast errors. *International Journal of Forecasting*, forthcoming.
Garcia, J.A., and Manzanares, A. (2007). Reporting biases and survey results—evidence from European professional forecasters. ECB Working Paper 836.
Gelain, P., Iskrev, N., Lansing, K.J., and Mendicino, C. (2019). Inflation dynamics and adaptive expectations in an estimated DSGE model. *Journal of Macroeconomics* 59, 258-277.
Gelfer, S. (2019). Data-rich DSGE model forecasts of the great recession and its recovery. *Review of Economic Dynamics* 32, 18-41.
Greenwood, R., and Shleifer, A. (2014). Expectations of returns and expected returns. *Review of Financial Studies* 27, 714-746.
Grothe, M., and Meyler, A. (2018). Inflation Forecasts: Are Market-Based and Survey-Based Measures Informative? *International Journal of Financial Research* 9, 171-188.
Harvey, D.I., and Newbold, P. (2003). The non-normality of some macroeconomic forecast errors. *International Journal of Forecasting* 19, 635–653.
Hubert, P., and Mirza, H. (2019). The role of forward- and backward-looking information for inflation expectations formation. *Journal of Forecasting*, forthcoming.
Keane, M. P., and Runkle, D. E. (1990). Testing the rationality of price forecasts: new evidence from panel data. *American Economic Review* 80, 714–735.
Koursaros, D. (2018). Learning expectations using multi-period forecasts. *Journal of Economics and Business*, forthcoming.
López Pérez, V. (2016a). Macroeconomic forecast uncertainty in the euro area. *Equilibrium* 11, 9-41.
López Pérez, V. (2016b). Does uncertainty affect non-response to the European Central Bank's survey of professional forecasters? *Economics: The Open-Access, Open-Assessment E-Journal* 10 (2016-25), 1–46.
Mankiw, N.G., and Reis, R. (2002). Sticky information versus sticky prices: a proposal to replace the New-Keynesian Phillips curve. *Quarterly Journal of Economics* 117, 1295–1328.
Morris, S., and Shin, H.S. (2002). Social Value of Public Information. *American Economic Review* 92, 1521–1534.
Nordhaus, W. (1987). Forecasting Efficiency: Concepts and Applications. *Review of Economics and Statistics*, 69, 667-674.
Poncela, P., Rodriguez, J., Sanchez-Mangas, R., and Senra, E. (2011). Forecast combination through dimension reduction techniques. *International Journal of Forecasting* 27, 224–237.
Ramos-Herrera, M.C., and Sosvilla-Rivero, S. (2018). Inflation, real economic growth and unemployment expectations: an empirical analysis based on the ECB Survey of Professional Forecasters, *Applied Economics*, forthcoming.
Reis, R. (2006). Inattentive consumers. *Journal of Monetary Economics* 53, 1761–1800.
Rossi, B., and Sekhposyan, T. (2016). Forecast rationality test in the presence of instabilities, with applications to Federal Reserve and survey forecasts. *Journal of Applied Econometrics* 31, 507-532.
Rossi, B., and Sekhposyan, T. (2018). Alternative tests for correct specification of conditional predictive densities. *Journal of Econometrics*, forthcoming.
Rubaszek, M., and Skrzypcznski, P. (2008). On the forecasting performance of a small-scale DSGE model. *International Journal of Forecasting* 24, 498–512.
Rudebusch, G.D., and Williams, J.C. (2009). Forecasting recessions: the puzzle of the enduring power of the yield curve. *Journal of Business & Economic Statistics* 27, 492-503.
Sims, C. (2003). Implications of rational inattention. *Journal of Monetary Economics* 50, 665–690.
Wang, Y.Y., and Lee, T.H. (2014). Asymmetric loss in the Greenbook and the Survey of Professional Forecasters. *International Journal of Forecasting* 30, 235–245.
Wieland, V., and Wolters M.H. (2011). The diversity of forecasts from macroeconomic models of the US economy. *Economic Theory* 47, 247–292.
Woodford, M. (2002). Imperfect common knowledge and the effect of monetary policy. In: Aghion, P., Frydman, R., Stiglitz, J., Woodford, M. (eds.), *Knowledge, Information and Expectations in Modern Macroeconomics*. Princeton University Press, 25–58.
Zhang, B. (2018). Real-time inflation forecast combination for time-varying coefficient models. *Journal of Forecasting*, forthcoming.

Annex I. List of contributors to the Funcas survey and date of first reply

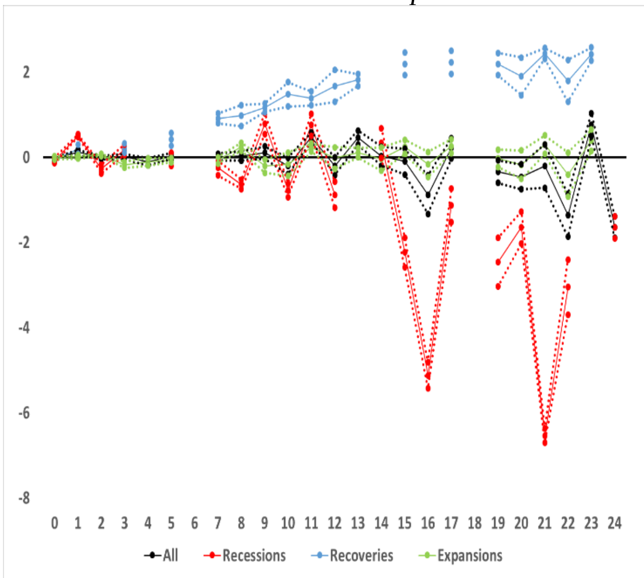
Name of the participating institution	*First reply*
Instituto de Estudios Económicos (IEE)	May 1999
Instituto de Crédito Oficial (ICO)	May 1999
Argentaria Joins BBVA in January 2000	May 1999
Banco Santander Central Hispano (BSCH) Banco Santander since September 2006	May 1999
La Caixa Caixabank since November 2015	May 1999
Funcas	May 1999
Banco Bilbao Vizcaya (BBV) BBVA since January 2000	May 1999
Instituto Complutense de Análisis Económico (ICAE)	May 1999
Centro de Predicción Económica (CEPREDE)	May 1999
Instituto Flores de Lemus de la Universidad Carlos III	May 1999
Caja Madrid Bankia since June 2011	June 1999
Analistas Financieros Internacionales (AFI)	November 1999
Caixa Catalunya Acquired by BBVA in May 2015	November 1999
Intermoney	November 1999
Consejo Superior de Cámaras de Comercio Cámara de Comercio de España since May 2017	March 2002
Cemex	February 2009
Repsol	April 2009
Centro de Estudios de Economía de Madrid (CEEM-URJC)	November 2009
Solchaga Recio & asociados	June 2010
Esade	November 2010
CEOE	September 2011
Instituto de Macroeconomía y Finanzas - Universidad CJC	September 2012
Axesor	May 2016

Annex II. Dates when the Funcas survey has been conducted

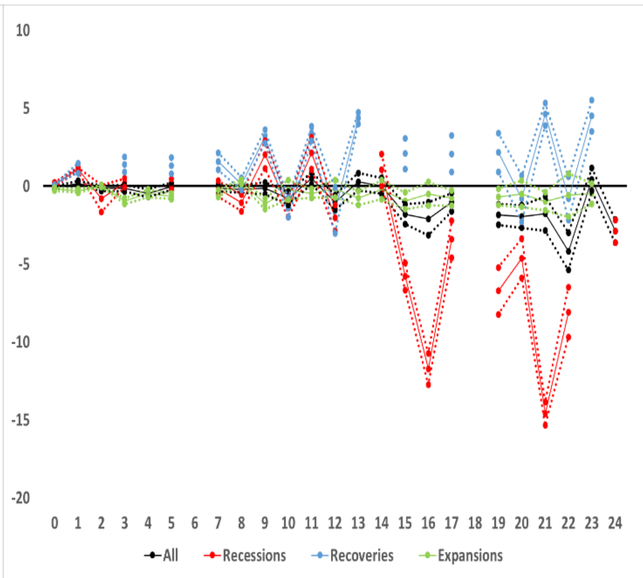
Year	Month											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1999					X	X	X		X	X	X	
2000	X		X		X		X		X		X	
2001	X		X		X		X		X		X	
2002	X		X		X		X		X			X
2003		X	X			X	X			X		X
2004		X		X		X	X			X		X
2005		X		X		X	X			X		X
2006		X		X		X	X			X		X
2007		X		X		X	X			X		X
2008		X		X	X		X			X	X	
2009		X		X		X	X		X		X	
2010		X		X		X	X		X		X	
2011		X		X		X	X		X		X	
2012		X		X		X	X		X			X
2013		X	X		X		X		X		X	
2014	X		X		X		X		X		X	
2015	X		X		X		X		X		X	
2016	X		X		X		X		X		X	
2017	X		X		X		X		X		X	

Annex III. Average forecast errors by horizon lengths and phase of the business cycle

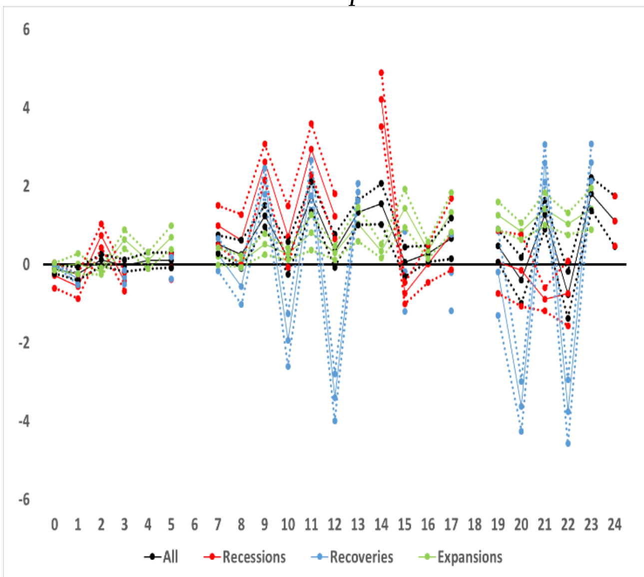
Private consumption



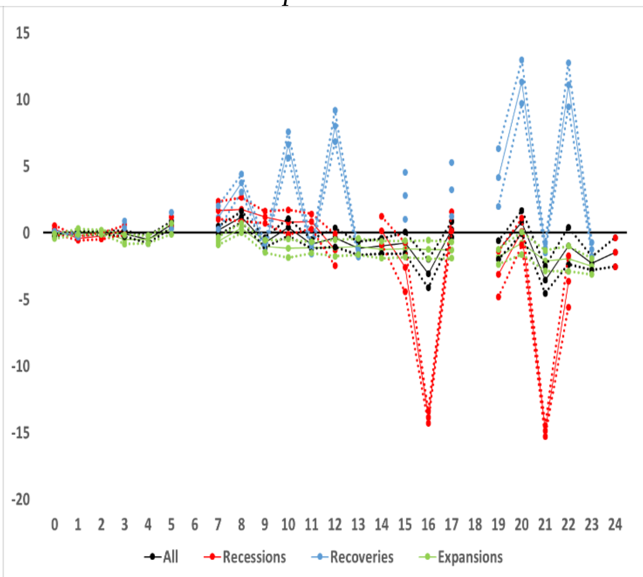
Private investment



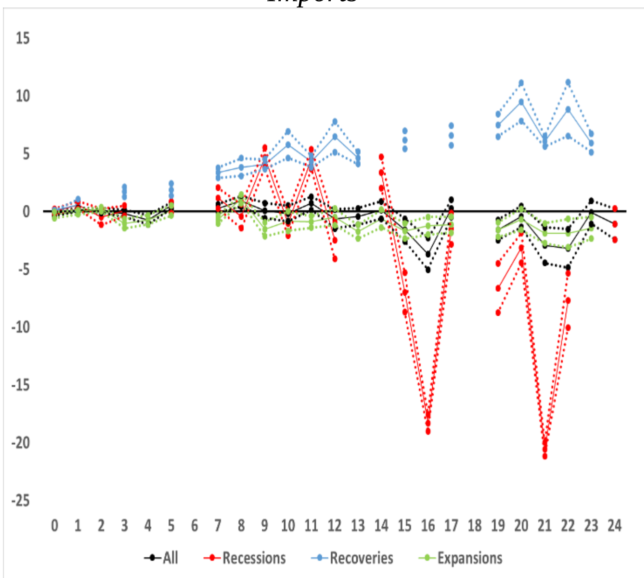
Government expenditure



Exports

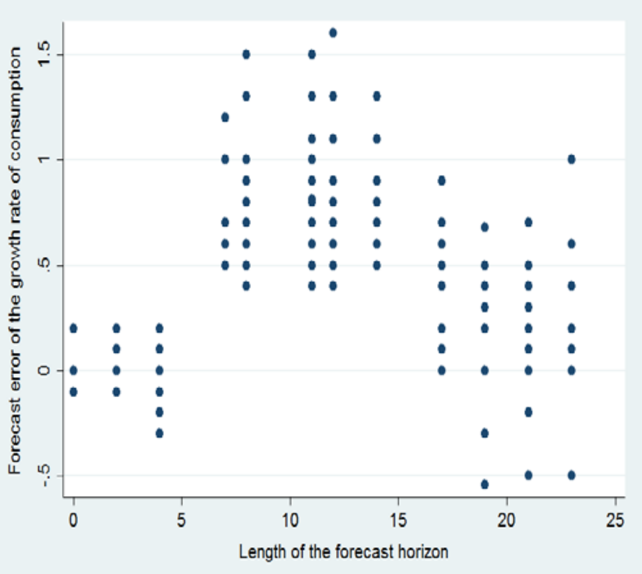


Imports

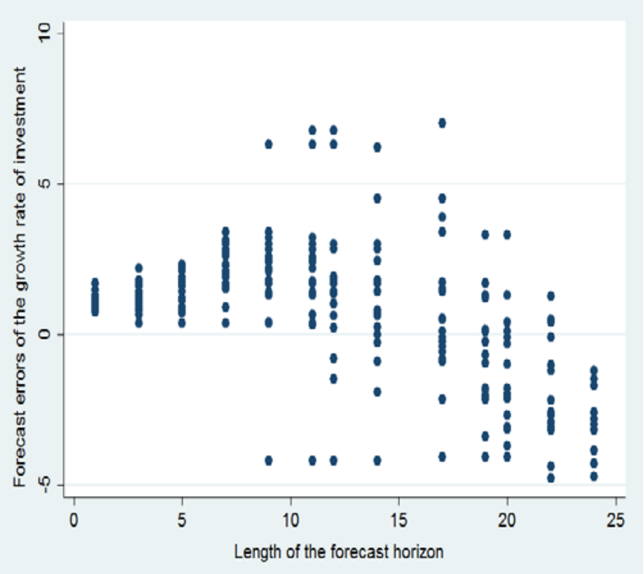


Annex IV. Individual forecast errors for subcomponents of GDP for selected target years

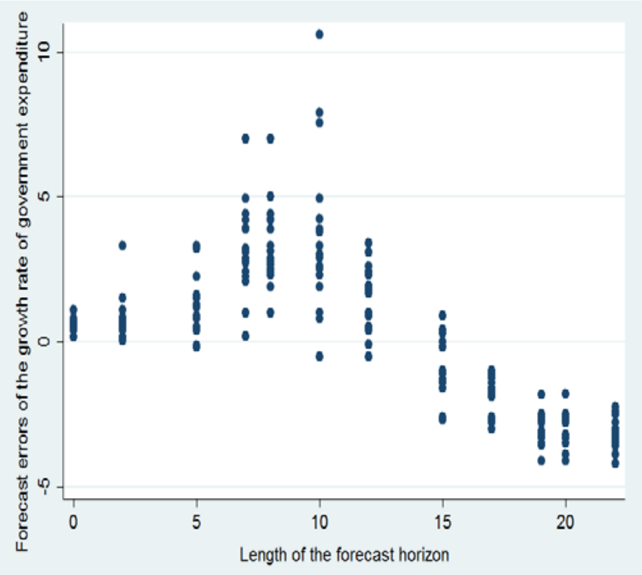
Private consumption 2003



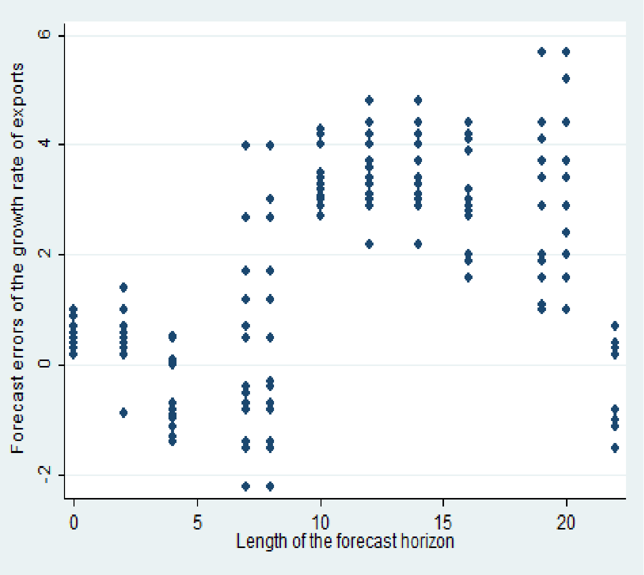
Investment 2013



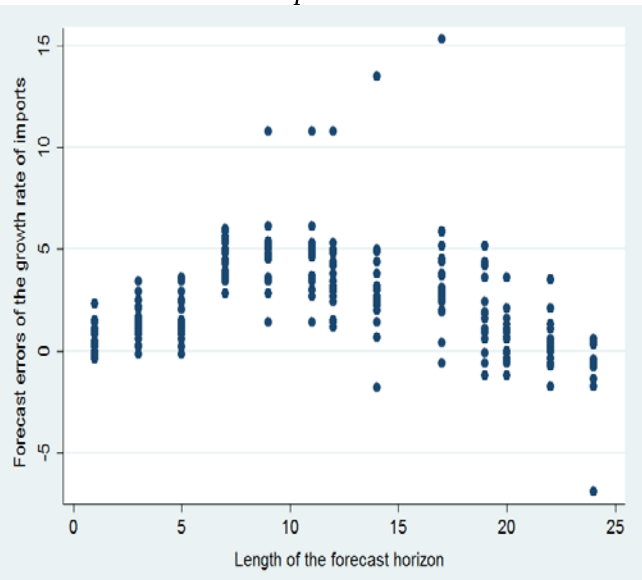
Government expenditure 2012



Exports 2006



Imports 2013



Annex V: Estimation results of the empirical model of forecast errors for the subcomponents of GDP

Private consumption

	M1	M2	M3	M4	M5	M6	M7	M8
β_1 (recession)		-1.436 (0.00)	-1.163 (0.00)	-3.536 (0.00)	0.632 (0.00)	-0.427 (0.00)	-0.523 (0.00)	-0.838 (0.00)
β_2 (recovery)			1.239 (0.00)	-0.976 (0.00)	0.335 (0.00)	-0.464 (0.00)	0.219 (0.19)	-0.363 (0.00)
δ_1 (recession started)				2.897 (0.00)	-0.221 (0.02)	-0.474 (0.00)	-0.093 (0.46)	-0.786 (0.00)
δ_2 (recovery started)				2.207 (0.00)	-0.268 (0.01)	-1.199 (0.00)	-1.067 (0.00)	-1.493 (0.00)
γ_1					-0.014 (0.00)	-0.016 (0.00)	-0.012 (0.00)	-0.004 (0.66)
γ_1 squared								0.000 (0.33)
γ_2 (recession)					-0.258 (0.00)	-0.218 (0.00)	-0.203 (0.00)	-0.164 (0.00)
γ_2 squared								-0.002 (0.24)
γ_3 (recovery)					-0.038 (0.00)	-0.043 (0.00)	-0.046 (0.00)	-0.159 (0.00)
γ_3 squared								0.005 (0.01)
γ_4 (recession started)					0.16 (0.00)	0.154 (0.00)	0.145 (0.00)	0.248 (0.00)
γ_4 squared								-0.004 (0.03)
γ_5 (recovery started)					0.151 (0.00)	0.094 (0.00)	0.095 (0.00)	0.249 (0.00)
γ_5 squared								-0.008 (0.00)
λ_1						0.094 (0.00)		0.105 (0.00)
λ_2						-0.046 (0.14)		-0.079 (0.01)
λ_3						-0.206 (0.00)		-0.259 (0.00)
λ_4						-0.307 (0.00)		-0.346 (0.00)
λ_5						-0.335 (0.00)		-0.358 (0.00)
λ_6						-0.388 (0.00)		-0.369 (0.00)
λ_7						-0.118 (0.00)		-0.091 (0.00)
λ_8						0.093 (0.00)		0.086 (0.00)
$\Sigma\lambda_i$ (expansion)							-0.685 (0.00)	
$\Sigma\lambda_i$ (recession)							-1.821 (0.00)	
$\Sigma\lambda_i$ (recovery)							-0.305 (0.08)	
α	-0.08 (0.00)							
Observations	3051	3051	3051	3034	3034	2806	2806	2806
Adj. R^2	0.00	0.20	0.29	0.47	0.61	0.68	0.71	0.70

Notes: P-values in parenthesis. Models M2 to M8 include individual dummy variables.

Investment

	M1	M2	M3	M4	M5	M6	M7	M8
β_1 (recession)		-2.722	-2.310	-7.601	2.143	2.303	-4.115	1.298
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
β_2 (recovery)			1.872	-3.258	1.729	0.820	-0.657	0.388
			(0.00)	(0.00)	(0.00)	(0.00)	(0.16)	(0.23)
δ_1 (recession started)				6.405	-0.689	-0.356	1.667	-0.492
				(0.00)	(0.00)	(0.22)	(0.00)	(0.11)
δ_2 (recovery started)				5.383	-0.537	-1.643	-0.146	-1.54
				(0.00)	(0.06)	(0.00)	(0.74)	(0.00)
γ_1					-0.067	-0.061	-0.067	-0.018
					(0.00)	(0.00)	(0.00)	(0.47)
γ_1 squared								-0.002
								(0.15)
γ_2 (recession)					-0.639	-0.669	-0.396	-0.363
					(0.00)	(0.00)	(0.00)	(0.00)
γ_2 squared								-0.014
								(0.00)
γ_3 (recovery)					-0.279	-0.239	0.147	-0.148
					(0.00)	(0.00)	(0.00)	(0.14)
γ_3 squared								-0.004
								(0.40)
γ_4 (recession started)					0.408	0.418	0.222	0.381
					(0.00)	(0.00)	(0.00)	(0.00)
γ_4 squared								0.002
								(0.60)
γ_5 (recovery started)					0.357	0.323	0.101	0.333
					(0.00)	(0.00)	(0.00)	(0.00)
γ_5 squared								-0.001
								(0.75)
λ_1						0.132		0.148
						(0.00)		(0.00)
λ_2						0.375		0.375
						(0.00)		(0.00)
λ_3						0.144		0.118
						(0.00)		(0.00)
λ_4						-0.305		-0.315
						(0.00)		(0.00)
λ_5						-0.293		-0.289
						(0.00)		(0.00)
λ_6						-0.455		-0.462
						(0.00)		(0.00)
λ_7						-0.002		-0.004
						(0.95)		(0.89)
λ_8						0.369		0.385
						(0.00)		(0.00)
$\Sigma\lambda_i$ (expansion)							-0.858	
							(0.00)	
$\Sigma\lambda_i$ (recession)							-1.339	
							(0.00)	
$\Sigma\lambda_i$ (recovery)							1.417	
							(0.00)	
α	-0.741							
	(-0.00)							
Observations	3051	3051	3051	3034	3034	2806	2806	2806
Adj. R^2	0.00	0.17	0.20	0.36	0.50	0.54	0.64	0.55

Notes: P-values in parenthesis. Models M2 to M8 include individual dummy variables.

Government expenditure

	M1	M2	M3	M4	M5	M6	M7	M8
β_1 (recession)		-0.007 (0.94)	-0.251 (0.01)	-1.576 (0.00)	-0.470 (0.01)	-1.930 (0.00)	-1.834 (0.00)	-3.280 (0.00)
β_2 (recovery)			-0.739 (0.00)	-1.962 (0.00)	-0.064 (0.67)	-1.623 (0.00)	-0.774 (0.00)	-2.133 (0.00)
δ_1 (recession started)				1.672 (0.00)	0.675 (0.00)	0.443 (0.02)	0.978 (0.00)	0.494 (0.01)
δ_2 (recovery started)				1.244 (0.00)	0.219 (0.26)	-0.074 (0.74)	0.924 (0.00)	0.695 (0.02)
γ_1					0.037 (0.00)	0.017 (0.00)	0.028 (0.00)	0.023 (0.16)
γ_1 squared								-0.000 (0.71)
γ_2 (recession)					-0.088 (0.00)	0.011 (0.42)	-0.067 (0.00)	0.325 (0.00)
γ_2 squared								-0.013 (0.00)
γ_3 (recovery)					-0.159 (0.00)	-0.004 (0.79)	-0.114 (0.00)	0.081 (0.18)
γ_3 squared								-0.003 (0.40)
γ_4 (recession started)					0.085 (0.00)	-0.030 (0.04)	0.038 (0.02)	-0.027 (0.64)
γ_4 squared								-0.001 (0.82)
γ_5 (recovery started)					0.058 (0.01)	-0.004 (0.78)	-0.054 (0.00)	-0.120 (0.04)
γ_5 squared								0.004 (0.15)
λ_1						-0.228 (0.00)		-0.256 (0.00)
λ_2						-0.470 (0.00)		-0.491 (0.00)
λ_3						-0.553 (0.00)		-0.527 (0.00)
λ_4						-0.386 (0.00)		-0.386 (0.00)
λ_5						-0.228 (0.00)		-0.193 (0.00)
λ_6						0.029 (0.49)		0.060 (0.13)
λ_7						0.107 (0.01)		0.103 (0.01)
λ_8						0.014 (0.73)		-0.053 (0.21)
$\Sigma\lambda_i$ (expansion)							-1.012 (0.00)	
$\Sigma\lambda_i$ (recession)							-0.785 (0.00)	
$\Sigma\lambda_i$ (recovery)							-2.838 (0.00)	
α	0.399 (0.00)							
Observations	2078	2078	2078	2061	2061	2041	2041	2041
Adj. R^2	0.00	0.06	0.07	0.12	0.14	0.47	0.56	0.49

Notes: P-values in parenthesis. Models M2 to M8 include individual dummy variables.

Exports

	M1	M2	M3	M4	M5	M6	M7	M8
β_1 (recession)		-0.626 (0.00)	0.039 (0.86)	-5.25 (0.00)	1.147 (0.03)	0.616 (0.18)	0.558 (0.33)	1.547 (0.00)
β_2 (recovery)			3.021 (0.00)	-1.999 (0.00)	0.173 (0.68)	0.336 (0.35)	2.078 (0.00)	0.771 (0.00)
δ_1 (recession started)				6.447 (0.00)	-0.146 (0.77)	-0.452 (0.28)	-0.202 (0.70)	-1.722 (0.00)
δ_2 (recovery started)				5.134 (0.00)	-0.424 (0.41)	-1.021 (0.03)	-0.200 (0.74)	-3.600 (0.00)
γ_1					-0.144 (0.00)	-0.147 (0.00)	-0.141 (0.00)	-0.123 (0.00)
γ_1 squared								-0.001 (0.46)
γ_2 (recession)					-0.370 (0.00)	-0.417 (0.00)	-0.496 (0.00)	-0.598 (0.00)
γ_2 squared								0.006 (0.46)
γ_3 (recovery)					-0.061 (0.19)	-0.303 (0.00)	-0.466 (0.00)	-0.601 (0.00)
γ_3 squared								0.012 (0.10)
γ_4 (recession started)					0.391 (0.00)	0.554 (0.00)	0.569 (0.00)	0.847 (0.00)
γ_4 squared								-0.010 (0.17)
γ_5 (recovery started)					0.379 (0.00)	0.358 (0.00)	0.295 (0.00)	1.180 (0.00)
γ_5 squared								-0.036 (0.00)
λ_1						-0.275 (0.00)		-0.277 (0.00)
λ_2						-0.136 (0.00)		-0.147 (0.00)
λ_3						-0.227 (0.00)		-0.245 (0.00)
λ_4						-0.334 (0.00)		-0.351 (0.00)
λ_5						-0.356 (0.00)		-0.353 (0.00)
λ_6						-0.189 (0.00)		-0.192 (0.00)
λ_7						-0.205 (0.00)		-0.223 (0.00)
λ_8						-0.070 (0.00)		-0.076 (0.00)
$\Sigma\lambda_i$ (expansion)							-1.843 (0.00)	
$\Sigma\lambda_i$ (recession)							-1.216 (0.00)	
$\Sigma\lambda_i$ (recovery)							-2.572 (0.00)	
α	-0.446 (0.00)							
Observations	3050	3050	3050	3033	3033	2805	2805	2805
Adj. R^2	0.00	0.03	0.10	0.23	0.30	0.56	0.63	0.57

Notes: P-values in parenthesis. Models M2 to M8 include individual dummy variables.

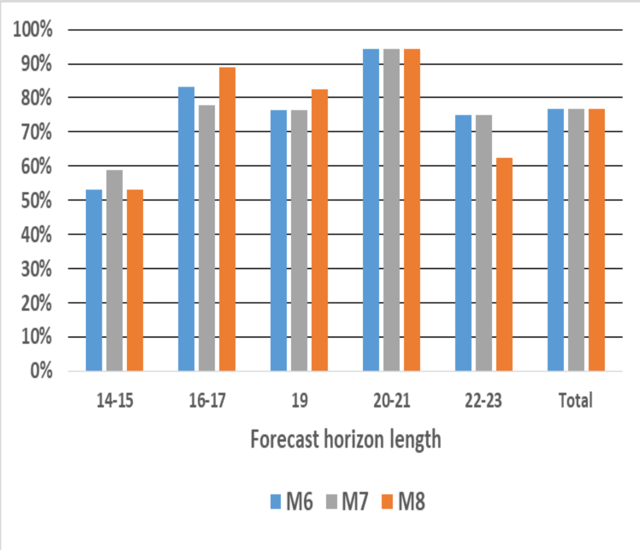
Imports

	M1	M2	M3	M4	M5	M6	M7	M8
β_1 (recession)		-2.814 (0.00)	-1.670 (0.00)	-9.879 (0.00)	1.548 (0.01)	0.037 (0.95)	-2.432 (0.01)	0.454 (0.27)
β_2 (recovery)			5.197 (0.00)	-2.255 (0.00)	1.708 (0.00)	1.961 (0.00)	3.026 (0.00)	1.916 (0.00)
δ_1 (recession started)				10.083 (0.00)	-0.499 (0.36)	-1.728 (0.00)	1.070 (0.26)	-3.630 (0.00)
δ_2 (recovery started)				7.126 (0.00)	0.077 (0.89)	-1.937 (0.00)	0.977 (0.33)	-4.966 (0.00)
γ_1					-0.135 (0.00)	-0.118 (0.00)	-0.109 (0.00)	-0.122 (0.00)
γ_1 squared								0.000 (0.83)
γ_2 (recession)					-0.715 (0.00)	-0.747 (0.00)	-0.671 (0.00)	-0.855 (0.00)
γ_2 squared								0.003 (0.70)
γ_3 (recovery)					-0.189 (0.00)	-0.423 (0.00)	-0.470 (0.00)	-0.650 (0.00)
γ_3 squared								0.009 (0.20)
γ_4 (recession started)					0.645 (0.00)	0.756 (0.00)	0.687 (0.00)	1.205 (0.00)
γ_4 squared								-0.017 (0.01)
γ_5 (recovery started)					0.456 (0.00)	0.424 (0.00)	0.192 (0.00)	1.375 (0.00)
γ_5 squared								-0.042 (0.00)
λ_1						-0.069 (0.00)		-0.077 (0.00)
λ_2						-0.077 (0.00)		-0.086 (0.00)
λ_3						-0.122 (0.00)		-0.152 (0.00)
λ_4						-0.300 (0.00)		-0.322 (0.00)
λ_5						-0.315 (0.00)		-0.310 (0.00)
λ_6						-0.160 (0.00)		-0.165 (0.00)
λ_7						-0.070 (0.00)		-0.092 (0.00)
λ_8						0.104 (0.00)		0.096 (0.00)
$\Sigma\lambda_i$ (expansion)							-1.047 (0.00)	
$\Sigma\lambda_i$ (recession)							-1.075 (0.00)	
$\Sigma\lambda_i$ (recovery)							0.053 (0.68)	
α	-0.486 (0.00)							
Observations	3050	3050	3050	3033	3033	2805	2805	2805
Adj. R^2	0.00	0.09	0.21	0.41	0.51	0.61	0.66	0.62

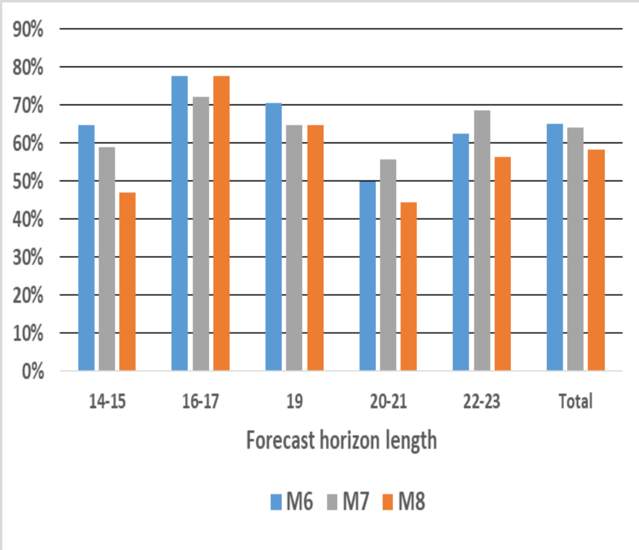
Notes: P-values in parenthesis. Models M2 to M8 include individual dummy variables.

Annex VI. Percentage of times each extended model outperforms the original model when predicting the consensus forecast errors of the subcomponents of GDP

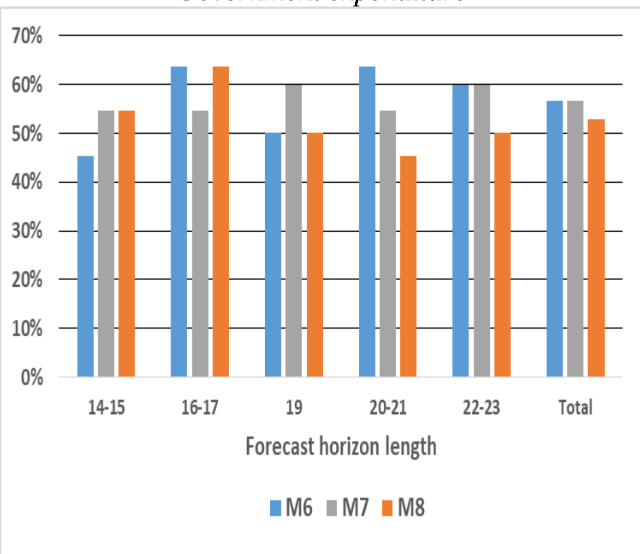
Private consumption



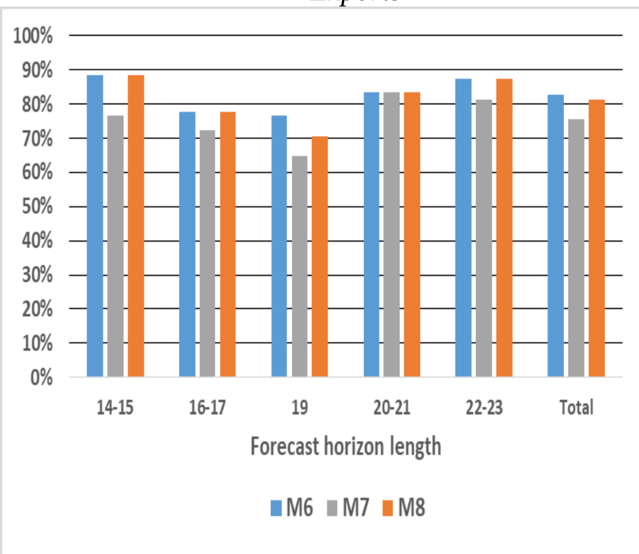
Investment



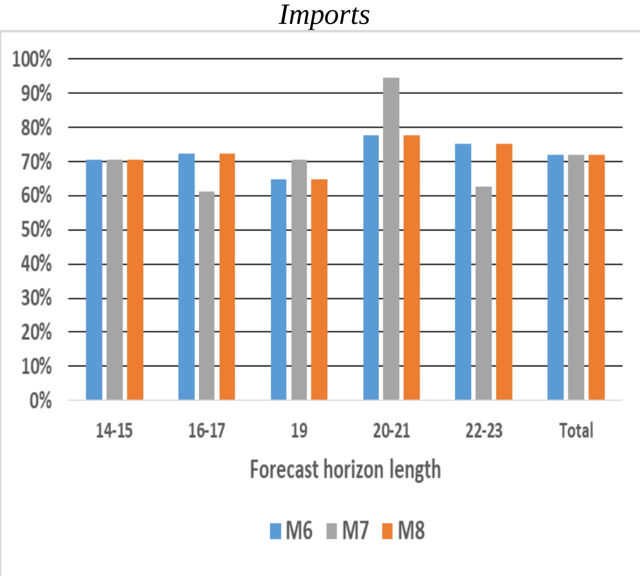
Government expenditure



Exports

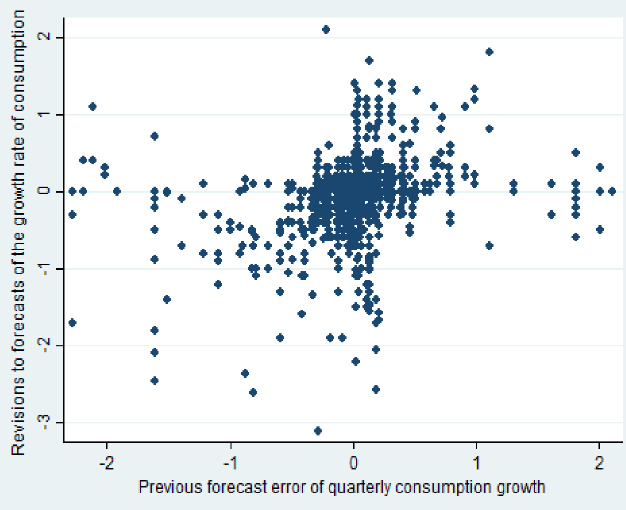


Imports



Annex VII. Relationship between forecast errors of the quarterly growth rate of private consumption and revisions to consumption forecasts

Year-on-year forecasts



Quarter-on-quarter forecasts