ANALYSIS

Bank of Finland staff forecasts: an evaluation

Finnish economy | 06.07.2018 | Eleonora Granziera, Juha Kilponen

AUTHORS

Eleonora Granziera

Juha Kilponen
Head of Monetary Policy and Research

Monetary policy decisions are based on assessment of the current and future state of the economy. In order to obtain forecasts, central banks build models, which are simplified representations of the complex interactions among macroeconomic variables. The Bank of Finland regularly publishes its forecasts, using a large set of data regarding current economic developments. Analysis of this large set of data includes the use of formal macroeconomic models, which are also employed to make projections for the future course of the economy. These projections represent the most likely values for the main macroeconomic variables of the Finnish economy.

This article documents the performance of Bank of Finland forecasts for GDP, inflation, unemployment and the components of GDP over the years 2004–2017. This period has been particularly challenging for forecasting. The financial crisis originating in the United States in 2007 spread globally. As a consequence Finland, as well as many other countries, experienced a severe contraction in output and a slow recovery. New policy measures were implemented in response to the crisis, and their effects on the economy were highly uncertain and difficult to anticipate.

Macroeconomic models and their ability to predict developments in the economy were called into question, as the models used for forecasting by Central Banks failed to predict the extent and duration of the crisis as well as the subsequent slow recovery. This led many monetary authorities to revise their models in order to incorporate more detailed description of the financial sector. In fact, recent analysis conducted at the Federal Reserve and European Central Bank suggests that models which include a richer set of financial variables would have been more successful in predicting the sharp decline of output in 2009 and the subsequent slow recovery.[1]

The Bank of Finland's main forecasting model, named AINO 2.0, is a stylized representation of the Finnish economy and describes complex relationships across a wide range of macroeconomic variables. The model was first used for the production of forecasts in 2004 and was updated in 2009 and again in 2015. The latest update was motivated by the need to explicitly model the

banking sector and its interaction with the real sector subsequent to the financial crisis.[2]

A previous evaluation of the Bank of Finland forecasts was conducted in 2011, shortly after the second update of the model, over a rather short sample spanning 2004–2010.[3] With an additional 7 years of data we are in a better position to examine the performance of the Bank of Finland staff forecasts using the latest vintage of the model.

The Bank's Monetary Policy and Research department produces forecasts twice a year, in June and December, for the current calendar year and the following two years. The forecasts are produced in conjunction with the Broad Macroeconomic Projection Exercise of the European System of Central Banks.

The forecasts are generated on the basis of the latest quarterly national accounts data and a set of assumptions regarding external developments, such as the future paths of exchange rates, foreign demand, oil and raw material prices plus market expectations of short-term interest rates.[4] Expert judgement on current economic conditions and likely future shocks hitting the economy are also incorporated as assumptions in the model. Judgement regarding current economic conditions reflects the information that can be obtained from the most recent short-term indicators but is not yet incorporated into the quarterly national accounts. In fact, statistical agencies release data pertaining to a month or quarter with delays. For example, GDP quarterly numbers are made available only two months after the reference quarter; e.g. the 2018Q1 first release of GDP was published at the end of May 2018. To overcome this issue, the Bank of Finland uses a set of nowcasting models [see also https://www.suomenpankki.fi/en/research/forecasting-models/]. Judgement regarding the future course of the economy reflects anticipated or highly likely events that are not explicitly modelled in AINO 2.0 but can significantly, if only temporarily affect aggregate demand or production, such as changes in fiscal policies, persistent sectoral shifts in allocation of the economy's resources, or changes in the Finnish export shares on world markets.

Another challenge to forecasting is that the data are subject to revision: e.g. Statistics Finland publishes revised values for past months or quarters. Each of these revised values is called a vintage. Some variables (e.g. GDP, exports and imports) are subject to larger revisions than others (e.g. HICP). In this evaluation exercise, forecasts are compared with the latest vintage available at the time the evaluation was made, i.e. the numbers published by Statistics Finland in March 2018. The final release of data should represent the 'true' value of the series.

Accuracy and unbiasedness of the forecasts

The forecasts are evaluated for accuracy and lack of bias, which are computed from the forecast

errors, i.e. the difference between the realized final values and the real time forecast. To measure bias we use the Mean Error (ME), simply the average of the forecast errors over the evaluation sample. If the forecast errors are on average close to zero, then the forecasts are said to be unbiased. This is a desirable property because it implies that the forecasters are not repeating the same mistake systematically. However, if the forecasts are consistently below (or above) the realizations, then the forecasts are negatively (or positively) biased. Note that an ME close to zero is also consistent with large forecast errors, as long as they are opposite in sign, so this statistic is not enough to judge the performance of the forecast. For this reason, we also compute three forecast accuracy measures: the Mean Absolute Error (MAE), the Root Mean Squared Forecast Error (RMSFE) and the Mean Change of Direction (MCD). As the names suggest, the MAE is the simple average of the absolute value of the forecast errors and the Root Mean Squared Forecast Error (RMSFE) is the squared root of the average of the squared forecast errors. The smaller the MAE or the RMSFE, the closer the forecasts are to the actual values, and therefore the more accurate they are. Both these measures of accuracy give equal weight to positive and negative errors of the same size, but the RMSFE penalizes errors of large magnitude more than the MAE does. The MCD is defined as the proportion of times a drop or increase in the series was correctly predicted. Trivially, more accurate forecasts translate into a higher MCD, with a maximum value of 100%. This statistic rewards correct predictions of the sign of the change in the series without giving consideration to the magnitude of the forecast error.

Figures 1 through 3 plot the variables[5] along with the forecasts for the current year and up to two years ahead. We distinguish between forecasts produced in June (in blue) and in December (in green). To highlight the importance of data revisions we also show the range of values taken by the variables throughout the different vintages (shaded gray area).

For most of the evaluation sample, GDP growth remained well below the 4% average registered in the previous decade, reaching a drop of about 8% in 2009. Although it experienced a quick rebound, output displayed another fall, less marked but more prolonged, in 2012–2014. Over these years the forecasts consistently predicted a much swifter recovery. In retrospect, this reflected the difficulty of the model and the economists to acknowledge a prolonged slowdown of productivity growth which was associated to the structural weakness of the Finnish economy and the loss of export market shares. Growth was more robust again in 2016–2017, and there the forecasts were less biased and more accurate.

The behaviour of the HICP resembles that observed in other euro area countries and exhibits the so-called twin puzzle: missing deflation in years 2009–2011, when inflation was around or even above target despite the prolonged and severe recession; missing inflation in years 2014–2017, when inflation was expected to rise thanks to the ongoing recovery. Despite these baffling

developments in HICP inflation, the Bank's forecasts tracked the inflation series quite closely, except for years 2011–2012 and 2015–2016, when the model predicted inflation would quickly revert to target, as output was also expected to grow faster. The puzzling behaviour of inflation can partially be explained by unexpected shocks in oil prices, which surged in 2011–2012 during the Arab Spring, only to plummet in 2014–2015.

The unemployment rate was steadily declining in the decade preceding the crisis and suddenly grew in 2009 as the recession hit the economy. Although the Bank of Finland correctly forecast a rise in the rate, it substantially over-predicted the size of the increase. This is because the Bank expected unemployment to behave in a similar fashion as in the crisis of the early 1990s, during which it increased swiftly, reaching a peak of almost 18%. In 2012–2015 unemployment rose further, mirroring the decline in output. Note that during the same period the Bank was instead optimistic, predicting a decline in the unemployment rate, consistent with the (optimistically) expected recovery in output and increase in the inflation rate.[6]

Revisions in GDP have been quite substantial relative to inflation and unemployment. The largest downward revision occurred in 2006 and amounted to 1.45%, while the following year registered the largest upward revision (0.92%). For the unemployment rate, the only significant revisions were made for the 2013 value, which was initially revised upward by 0.3% and subsequently revised downward by the same amount. Revisions to inflation are negligible (up to 0.02%).

The descriptive statistics summarizing the evaluation results for GDP, inflation and unemployment are presented in Table 1.

In general, forecasts produced in December are less biased and more accurate than those produced in June, reflecting the larger data set available at the time the forecasts are made, such as a new release of National Accounts by Statistics Finland for quarterly GDP and new releases of monthly HICP, unemployment figures and short-term indicators. This finding is also clearly seen in Charts 1 through 3, in which the green lines are closer to the actual data than the blue lines.

Forecasts for GDP are overall negatively biased, i.e. the forecasts systematically over-predict output growth during the sample, and the bias increases (in absolute value) with the forecast horizon. The forecasts errors are usually negative (positive), i.e. we tend to over-predict (under-predict) GDP growth, when GDP growth is below (above) the mean. This reflects, in general, the model's tendency to converge towards its historical mean rates too fast.

Note that over the evaluation sample, output fell drastically in 2009 and fell again substantially in 2012–2013. In these years we observe by far the largest, negative forecast errors. For example, in 2007 the Bank of Finland forecast that output would grow by 2.5% in 2009, while it fell by 8%,

resulting in a forecast error of -10.5. Removing 2009 from the evaluation substantially improves the bias, which reduces to -0.05 for current year and 0.08 and -0.12 for one and two years ahead, respectively.

Both inflation and unemployment display a much smaller bias.[7] For inflation it is generally positive and larger at one and two year horizons. This is due to an over-prediction of the model, which forecasts that within about two years HICP inflation will return to target, i.e. close to but under 2%.

The bias for unemployment is positive at two years ahead and negative for shorter horizons. From Chart 3 it is clear that by far the largest forecast errors were made for the years 2009–2011. Because the errors were opposite in sign, e.g. 2.8% for 2010 and -3% for 2011 for the two years ahead horizon, they each other cancel out, and the resulting bias is therefore small.

In terms of forecasting performance for all variables and measures considered, accuracy deteriorates with the forecasting horizon. Forecasts are more accurate in December than in June. Output records the highest MAE and RMSFE, but their values are, again, reduced considerably if we exclude 2009 from our evaluation sample, dropping by about 40% for one-step-ahead and two-step-ahead predictions.

The results for the components of GDP are reported in Table 2. Imports, exports and private investment prove difficult to forecast, as they exhibit large bias and RMSFEs. The model tends to under-predict these components in the current year and over-predict two years ahead. However if we focus on the MCD, then the forecasting performance is comparable with the other components of GDP. Also, the large RMSFE associated with the exports, imports and investment series reflect the high volatility and large revisions of these variables. To appreciate the high volatility of the series, note that, when we scale the RMSFE by the standard deviation of the series computed over the evaluation sample, we obtained ratios that are close to those of the other variables. Finally, the bias and accuracy of the forecasts on import and exports crucially depend on the appropriateness of the assumptions regarding future developments in the exchange rate and foreign demand. For example, the large positive bias in the export forecasts for current year forecasts and the large negative bias for one and two years ahead mirror the large bias in the forecasts for external demand (the bias in external demand can amount up to 90% of the bias in exports). Similarly, the RMSFE for exports is comparable in magnitude to the RMSFE of external demand for all forecast horizons.

Forecast performance relative to benchmarks

Rather than focusing exclusively on absolute performance, it can be informative to also comment

on the relative forecasting performance of the AINO model compared with simple benchmarks. To this end, the accuracy of the Banks' forecasts is compared against the accuracy of three alternative forecasts: the real time mean, an autoregressive model of order one and forecasts produced by the Ministry of Finance. To measure relative forecast accuracy, we compute the relative mean squared error, defined as the ratio between the RMSFE from the Banks' forecasts and the RMSFE from one of the alternative forecasts. A value lower (higher) than one indicates a better (worse) performance of the forecasts from the Bank of Finland's model. Our results are reported in Table 3.

The real time mean is computed for every date in our sample as the average taken by the series of the previous eight years. This moving average, which discards values further back in time, is well suited to account for the turbulent economic conditions of our evaluation sample. For all variables and all forecast horizons, the Banks' forecasts are more accurate than the mean, although the relative RMSFE approaches one as the forecast horizon increases. This suggests that at longer horizons the forecasts from the AINO model converge to the real time mean of the series.

The autoregressive model of order one uses the most recent value to forecast future values. This simple model has been found to be a very competitive benchmark (Chauvet and Piger 2012). Generally, the AINO forecasts are more accurate than the forecasts from the autoregressive model, although gains are smaller than over the real time mean and tend to dissipate for forecasts for two years ahead.

The performance of the Bank's forecasts and Ministry of Finance forecasts are quite comparable.[8] Occasional differences can be attributed to outliers which greatly affect the RMSFE in such a small evaluation sample. For example, for unemployment, the Bank's forecasts appear more accurate than the Ministry of Finance forecasts for the current year. However, for two years ahead the Ministry of Finance forecasts display a somewhat lower RMSFE. It turns out that this last result is driven by the forecast of the 2015 unemployment rate made in December 2013, when the Bank was optimistic about the recovery of the labour market. For output, the performance is quite comparable, except for the June forecasts made for the current year. For this case the Ministry of Finance produced more accurate forecasts overall and the Bank made a larger forecast error in 2012, as it did not foresee the double-dip recession. The Bank's forecasts seem to be more accurate at all horizons for inflation. Note that while the Bank of Finland forecasts HICP inflation, the Ministry of Finance focuses on CPI inflation. These two series overlap for most of the sample, except in 2009, when CPI dropped to zero, while HICP remained closer to target (1.63%).

General equilibrium models have benefits that go well beyond forecast accuracy

We have documented the performance of the Bank of Finland forecasts over a turbulent period of time characterized by large and abrupt changes in the economic and policy environment. In general, the forecasts display small bias and good accuracy. However, sizable data revisions and large volatility can negatively impact bias and accuracy for some series. Furthermore, the properties of the forecasts depend on the appropriateness of expert judgement and assumptions on the international economy, which are external to the model. The Bank's forecasts compare well with those from simple but competitive benchmark models as well as with those produced by external institutions. Still, the short sample makes some of our results sensitive to outliers and does not allow for rigorous statistical testing on the measures of bias and accuracy used. Regardless, employing a general equilibrium model to produce forecasts provides benefits that go beyond forecast accuracy, such as internal consistency and the ability to provide a structural interpretation for the forecast, benefits which are not easily attainable with simple univariate models. Both properties derive from the fact that the AINO model's economic relations among macro-variables are largely derived from modern macroeconomic theory. Forecast values for different macroeconomic variables are then internally consistent because they satisfy these relations.[9] For similar reasons, models like AINO can be viewed as devices useful for storytelling as they can inform policymakers about future developments taking into account complex interactions of agents in the economy as well as future policy decisions.

References

Cai, M. – Del Negro, M. – Giannoni, M.P. – Gupta, A. – Li, P – Moszkowski, E. (2018) 'DSGE Forecasts of the Lost Recovery', NYFED wp. 844.

Chauvet, M. – Potter, S. (2012) 'Forecasting Output', in Handbook of Economic Forecasting, Volume 2, Elsevier.

Kilponen, J. – Orjasniemi, S. – Ripatti, A. – Verona, F. (2016) The Aino 2.0 model, Research Discussion Papers 16/2016, Bank of Finland.

Linde, J. – Smets, F. – Wouters, R. (2016) Challenges for Central Banks' Macro Models, Working Paper Series 323, Sveriges Riksbank.

Newby, E. – Orjasniemi, S. (2011), 'Bank of Finland's forecast errors in 2004–2010', Bank of Finland Bulletin 3/2011.

Chart 1.

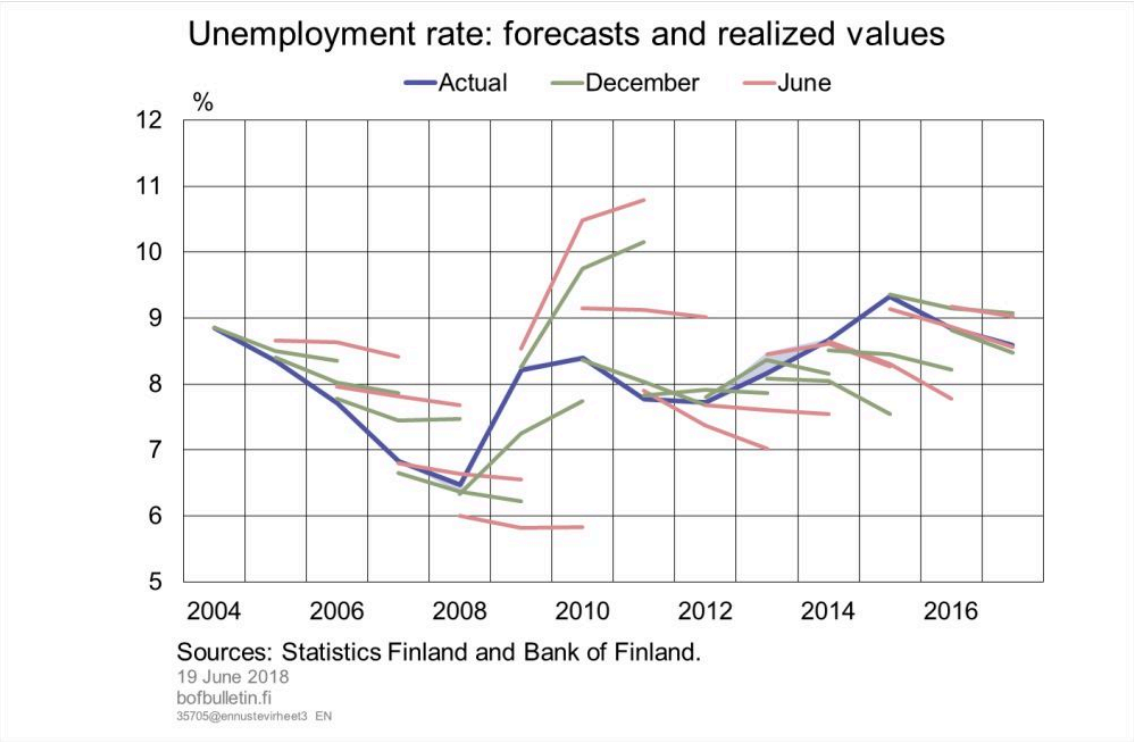
Unemployment rate: forecasts and realized values
Actual
December
June
%
12
11
10
9
8
7
6
5
2004
2006
2008
2010
2012
2014
2016
Sources: Statistics Finland and Bank of Finland.
19 June 2018
bofbulletin.fi
35705@ennustevirheet3 EN

Chart 2.

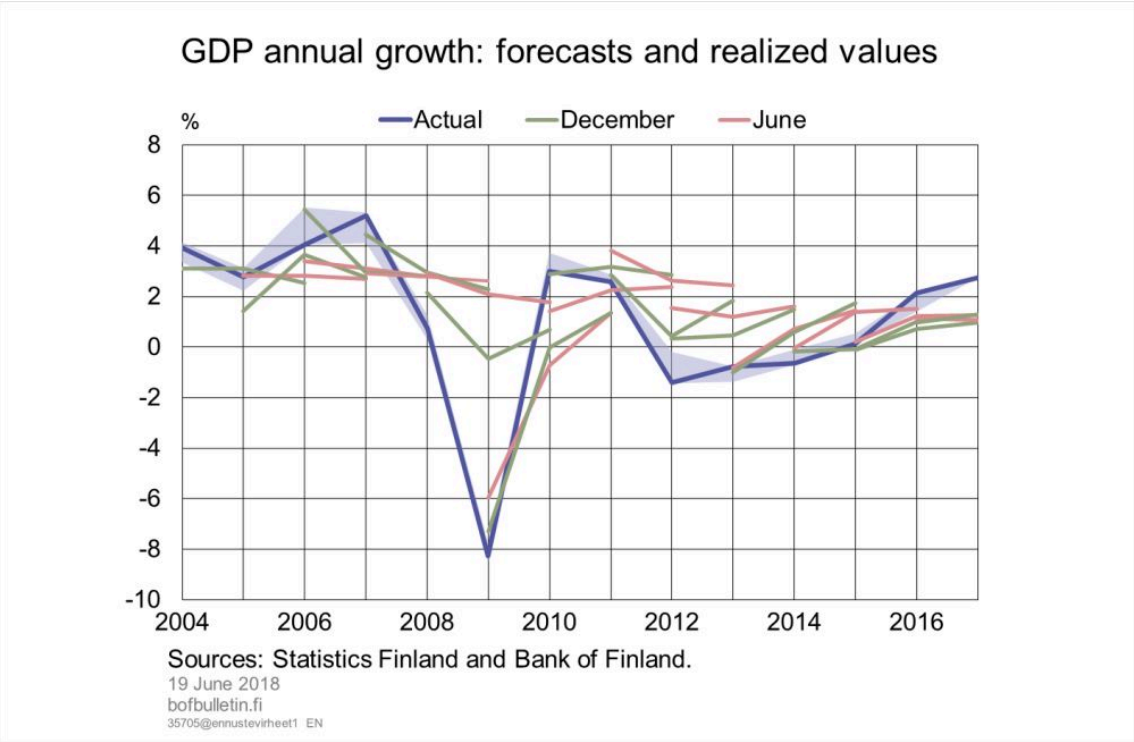
GDP annual growth: forecasts and realized values
%
Actual
December
June
8
6
4
2
0
-2
-4
-6
-8
-10
2004
2006
2008
2010
2012
2014
2016
Sources: Statistics Finland and Bank of Finland.
19 June 2018
bofbulletin.fi
35705@ennustevirheet1 EN

Chart 3.

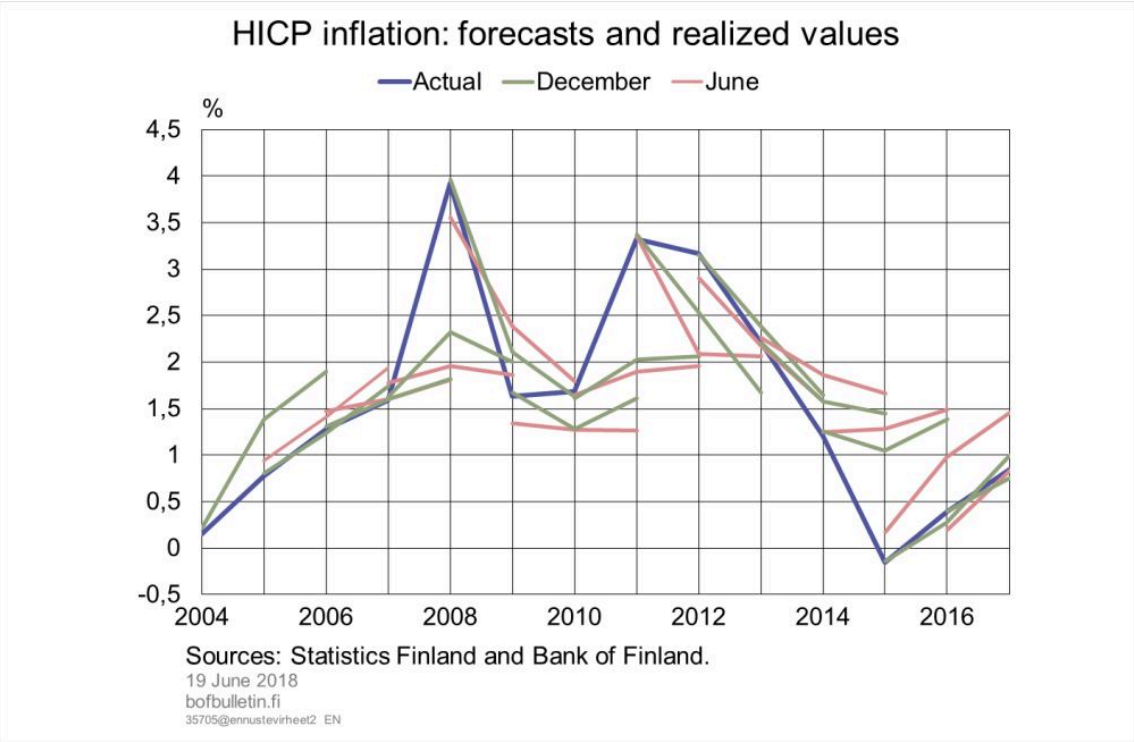

Table 1.

Bias and accuracy for GDP growth, HICP and unemployment, 2004–2017

	GDP			HICP			Unemployment rate		
	June	Dec	overall	June	Dec	overall	June	Dec	overall
ME current	-0.25	-0.14	-0.19	0.02	-0.02	0.00	-0.12	0.02	-0.05
ME 1 year ahead	-0.92	-0.50	-0.71	0.12	0.10	0.11	-0.13	-0.06	-0.09
ME 2 years ahead	-1.49	-1.07	-1.28	0.09	0.10	0.09	0.14	0.03	0.08

MAE current	1.20	0.81	1.00	0.17	0.03	0.10	0.24	0.07	0.16
MAE 1 year ahead	2.58	1.85	2.20	0.71	0.54	0.62	0.86	0.47	0.66
MAE 2 years ahead	2.78	2.80	2.79	0.92	0.82	0.87	1.43	0.95	1.19
RMSFE current	1.53	0.96	1.27	0.20	0.04	0.14	0.32	0.08	0.23
RMSFE 1 year ahead	3.64	2.65	3.16	0.94	0.73	0.84	1.13	0.60	0.91
RMSFE 2 years ahead	3.89	3.73	3.81	1.18	1.05	1.11	1.62	1.18	1.42
MCD current	0.67	0.85	0.76	1.00	0.92	0.96	0.92	1.00	0.96
MCD 1 year ahead	0.36	0.67	0.52	0.45	0.67	0.56	0.64	0.83	0.73
MCD 2 years ahead	0.40	0.27	0.34	0.10	0.36	0.23	0.50	0.55	0.52
Rel RMSFE current	0.45	0.28	0.37	0.17	0.03	0.12	0.40	0.11	0.29
Rel RMSFE 1 year	1.04	0.75	0.90	0.77	0.60	0.69	1.38	0.73	1.09
Rel	1.11	1.06	1.08	0.97	0.86	0.92	1.98	1.44	1.72

RMSFE 2 years									

Table 2.

Bias and accuracy for components of GDP, 2004–2017

	IMPORTS			EXPORTS			OV. CONSUMPTION		
	June	**Dec**	**overall**	**June**	**Dec**	**overall**	**June**	**Dec**	**overall**
ME current	1.60	4.41	3.06	0.87	2.05	1.54	-0.13	0.26	0.08
ME 1 year ahead	-0.83	0.56	-0.11	-1.67	-0.55	-1.09	-0.03	0.04	0.01
ME 2 years ahead	-2.59	-1.89	-2.22	-3.54	-2.51	-3.00	-0.26	-0.09	-0.17
RMSFE current	2.74	5.48	4.38	3.25	3.44	3.41	0.97	0.56	0.78
RMSFE 1 year ahead	7.82	5.53	6.73	8.66	5.81	7.31	0.84	0.72	0.78
RMSFE 2 years ahead	7.43	7.00	7.21	8.59	8.22	8.40	1.04	0.80	0.92
MCD current	0.58	0.54	0.56	0.75	0.69	0.72	0.50	0.69	0.60
MCD 1 year ahead	0.64	0.58	0.61	0.45	0.58	0.52	0.55	0.75	0.65
MCD 2	0.70	0.45	0.58	0.40	0.45	0.43	0.40	0.73	0.56

years ahead									
Rel RMSFE current	0.40	0.79	0.63	0.42	0.44	0.43	1.20	0.69	0.97
Rel RMSFE 1 year	1.15	0.82	0.99	1.08	0.72	0.91	1.08	0.93	1.00
Rel RMSFE 2 years ahead	1.10	1.03	1.06	1.07	1.02	1.05	1.34	1.03	1.19
	INVESTMENT			**CONSUMPTION**			**GOV. INVESTMENT**		
	June	**Dec**	**overall**	**June**	**Dec**	**overall**	**June**	**Dec**	**overall**
ME current	0.26	-0.02	0.11	0.19	0.03	0.10	-0.97	0.06	-0.44
ME 1 year ahead	-1.54	-0.52	-1.01	0.04	0.27	0.16	-0.23	-0.38	-0.31
ME 2 years ahead	-2.87	-2.32	-2.58	-0.62	-0.13	-0.37	1.31	0.51	0.89
RMSFE current	3.05	2.50	2.78	0.68	0.78	0.74	4.45	3.62	4.04
RMSFE 1 year ahead	7.32	4.84	6.15	1.32	1.14	1.23	5.34	4.73	5.03
RMSFE 2 years ahead	6.70	6.49	6.59	1.38	1.37	1.38	4.94	5.32	5.14

MCD current	0.58	0.69	0.64	0.92	0.62	0.77	0.42	0.85	0.63
MCD 1 year ahead	0.36	0.58	0.47	0.55	0.50	0.52	0.36	0.42	0.39
MCD 2 years ahead	0.70	0.55	0.62	0.40	0.45	0.43	0.40	0.36	0.38
Rel RMSFE current	0.52	0.43	0.48	0.52	0.60	0.56	0.96	0.78	0.87
Rel RMSFE 1 year	1.21	0.80	1.02	1.02	0.89	0.95	1.23	1.09	1.16
Rel RMSFE 2 years ahead	1.11	1.08	1.09	1.07	1.06	1.07	1.14	1.22	1.18

Table 3.

Relative forecasting performance for GDP growth, HICP and unemployment, 2004–2017

	GDP						**HICP**					
	Mean		**AR**		**MoF**		**Mean**		**AR**		**Mo**	
	June	**Dec**	**June**	**Dec**	**June**	**Dec**	**June**	**Dec**	**June**	**Dec**	**June**	
current year	0.42	0.28	0.45	0.29	1.19	1.08	0.15	0.03	0.20	0.04	0.92	
one year ahead	0.93	0.70	0.99	0.75	1.00	1.02	0.66	0.53	0.80	0.63	0.74	

two years ahead	0.96	0.97	1.05	1.04	0.93	1.05	0.82	0.75	0.95	0.88	0.79
	UNEM										
	Mean		**AR**		**MoF**						
	June	**Dec**	**June**	**Dec**	**June**	**Dec**					
current year	0.23	0.06	0.50	0.14	0.78	0.59					
one year ahead	0.73	0.36	1.25	0.69	1.17	0.85					
two years ahead	0.98	0.63	1.70	1.25	1.10	1.35					

Footnotes

1. See Cai et al. (2018) and Lindé et al. (2016). ↑
2. Kilponen, Orjasniemi, Ripatti and Verona (2016) provide a detailed description of the latest version of the AINO model. ↑
3. See Newby and Orjasniemi (2011). ↑
4. The appropriateness of these assumptions might affect the accuracy of the forecasts. However, in this analysis we are not conducting a systematic evaluation of the impact of the external assumptions on the accuracy of the forecasts. ↑
5. For each year t, output growth and its components, as well as inflation are defined as: (xtQ1 + xtQ2 + xtQ3 + xtQ4)/ (xt-1Q1 + xt-1Q2 + xt-1Q3 + xt-1Q4) with xtQj the level of the series in quarter j of year t. Unemployment is the average of the monthly rates. ↑
6. Note that the unemployment rate is not part of the core theoretical model; it is forecast using a simple empirical model with feedback from the core model's labour market variables such as total hours worked. ↑
7. Because of the small size of our sample, we did not conduct any formal forecast evaluation test for the bias or relative forecast accuracy, as the results from these tests would be unreliable. ↑
8. The Ministry of Finance did not publish its forecasts in December 2007 and started to publish forecasts for two years ahead only in December 2009. To construct the relative

RMSFE we recomputed the RMSFE for the Bank's forecasts over the overlapping sample of observations. ↑

9. One example of these relations is the Phillips curve. Then, based on this relationship between inflation and economic activity, the model would predict an increase in inflation associated with an increase in demand and positive output growth. ↑

Key words

forecast error, gross national product, inflation, revisions, unemployment