International Journal of Forecasting 32 (2016) 23–33

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Herding behavior of business cycle forecasters

Jan-Christoph Rülke [a,*], Maria Silgoner [b], Julia Wörz [b]

[a] EBS Universität für Wirtschaft und Recht, Department of Management and Economics, Rheingaustr. 1, 65375 Oestrich-Winkel, Germany

[b] Oesterreichische Nationalbank (OeNB), Otto-Wagner-Platz 3, 1090 Vienna, Austria

ARTICLE INFO

Keywords:
Anti-herding
Business cycle
Forecasting
Economic uncertainty
Herding

ABSTRACT

Using a large international data set, we analyze whether business cycle forecasters herd or anti-herd. In general, we find evidence for anti-herding, i.e. forecasters appear to scatter their forecasts deliberately away from the forecasts of others. Anti-herding tends to be more prevalent for the longer (next year) horizon. There is some evidence for a reduced level of anti-herding at times of increased forecast uncertainty and when the forecasts are being revised more substantially.

1. Introduction

Business cycle and growth expectations play a major role in understanding macroeconomic relationships. They also determine the extent to which economic policy agents, including central banks, can influence macroeconomic outcomes. One way to deal with forecast uncertainty is to pool the expectations of professional forecasters (Zarnowitz, 1984) in order to hedge against the errors of individual forecasters, thus improving the forecast quality. Such surveys of professional forecasters are provided by either central banks or private companies. The idea of these consensus forecasts is that, although individual forecasters may outperform the average of a group of forecasters in certain cases, an individual forecaster rarely outperforms others systematically. Zarnowitz and Lambros (1987) find that the forecast errors of consensus forecasts are smaller than those of most individual forecasters. Batchelor (2001) shows that consensus forecasts are more accurate than the projections published by the OECD or the IMF.

The reliability and superiority of consensus forecasts depends crucially on whether the forecasters actually reveal their own best forecast or behave strategically, i.e., show herding or anti-herding tendencies. Forecaster herding arises if the forecasters ignore their private information and instead follow the forecasts of others (Scharfstein & Stein, 1990). For example, Bewley and Fiebig (2002) show that interest rate forecasters tend to indicate values in the safe consensus range, in order to avoid sticking their neck out with "extreme" forecasts. This is because a poor forecast may not damage a forecaster's reputation if other forecasters also delivered poor forecasts. Thus, herding behavior biases the distribution towards the mean. Forecaster herding should not be confused with forecast clustering, where similar forecasts may be observed because all forecasters have access to the same set of economic data and similar forecast techniques. Herding behavior, on the other hand, refers to forecasters deliberately deviating from their best private forecasts for strategic reasons.

Forecaster anti-herding may arise if forecasters, for strategic or other reasons, deliberately scatter their forecasts away from the forecasts of others. This may arise when a forecaster's income (or reputation) depends not only on the accuracy of their own forecasts, but also on their relative performances. If some of the customers of professional forecasters buy forecasts only occasionally,

* Corresponding author.
E-mail address: jan-c.ruelke@whu.edu (J.-C. Rülke).

http://dx.doi.org/10.1016/j.ijforecast.2015.02.004
0169-2070/

and pick the forecasters with the best performances in the last period, forecasters have a strong incentive to differentiate their forecasts from those of others. An "extreme" forecast may have a small probability of being accurate, but the expected payoff can be high if the forecasters can attract new customers in the case of a stroke of luck (Laster, Bennett, & Geoum, 1999).

From the perspective of monetary and economic policy, a knowledge regarding the reliability of forecasts is of the utmost importance. Forecaster herding implies not only a smaller level of forecast heterogeneity, but also forecast inertia, so that the adjustment of the consensus forecast to newly available economic data is delayed. As forecasts may themselves influence economic reality through self-fulfilling prophecy effects (Grisse, 2009), it is essential to know whether pooled forecasts are subject to (anti-) herding biases or not.

There is a substantial body of literature on the analysis of herding behavior in the forecasting industry. For GDP growth forecasts of the Consensus Economics data set for the USA, the UK and Japan, Gallo, Granger, and Jeon (2002) find that forecasters pay too much attention to leaders in the group. Pons-Novell (2003) uses unemployment rate forecasts for the USA, published in the Livingston survey, to analyze whether professional forecasters behave strategically in order to maximize publicity, salary or their prestige. He finds that age and reputation effects matter. Lamont (2002) shows that, as forecasters become older and more established, they produce more radical forecasts, indicating reputation effects. Pierdzioch, Rülke, and Stadtmann (2010) use the test proposed by Bernhardt, Campello, and Kutsoati (2006) to investigate (anti-)herding behavior in the forecasting of financial variables (commodity prices and exchange rates), and find evidence of anti-herding. Freedman (2013) shows that the herding behavior of macroeconomic forecasters is related inversely to the sizes of past forecast errors, and decreases when other forecasters make large errors.

In this paper, we analyze whether herding or anti-herding behaviors are inherent in GDP growth forecasts. We are the first to analyze this topic over the period of the global economic and financial crisis. Our sample period allows us to investigate a novel hypothesis in this context which has been disregarded by the literature to date: the stage within the business cycle, and therefore the level of economic uncertainty, may influence researchers' incentives to either hide in the comfortable consensus middle or go out on a limb with extreme forecasts. More specifically, we investigate (anti-)herding tendencies during times of elevated forecast uncertainty, relative to normal times. We study forecaster herding using the empirical test developed by Bernhardt et al. (2006), and find evidence of anti-herding behavior for most industrial economies, but signs of forecaster herding for emerging economies. We relate this finding to the high incidence of economic and financial crises in these countries, since tests confirm that forecasters tend to herd in times of high forecast uncertainty and we do not find any statistically significant relationship between a country's stage of economic development and herding behavior.

2. Data

In our empirical analysis, we use monthly survey data on business cycle forecasts compiled by Consensus Economics, which has been publishing average forecasts for a broad set of countries since October 1989. Today, the monthly survey covers forecasts from more than 1000 economists worldwide for 75 countries and various variables (GDP growth, inflation, the current account balance, interest rates). However, disaggregated forecast data (i.e., the forecasts of individual survey participants) are only available for a subset of 45 countries. Thus, our analysis is limited to these 45 countries. All forecasters are located in the country for which they are forecasting, and hence, should have a very good idea concerning business cycle developments.[1]

The numbers of forecasters and forecasts vary across countries. While we have at our disposal about 800 forecasts submitted by a group of 18 forecasters for the Philippines, we can also study data from 68 forecasters who published more than 14,000 business cycle forecasts for the United Kingdom. Our sample period ends in December 2011, and includes a total of 226,851 business cycle forecasts published by 1604 forecasters. Among other indicators, Consensus Economics publishes forecasts each month for the average annual growth rates of GDP for the current and next year. Thus, there are 24 consecutive consensus forecasts for a given calendar year.[2]

There are at least three reasons why our data set is particularly suitable for studying the herding instinct of business cycle forecasters. First, because the poll is conducted during the first week of each month and released within the second week, it is a timely and frequent indicator for growth expectations. Second, the dataset has large cross-sectional and time series dimensions of more than twenty years.

Third, individual forecasts are published together with the name and affiliation of the forecaster. This enables us to evaluate the performance of an individual forecaster. Since the survey is non-anonymous, forecasters' performances can be expected to have an effect on their reputations. This link between performance and reputation may strengthen forecasters' incentives to herd, since a poor forecast may not damage a forecaster's reputation if the other forecasters also delivered poor forecasts. On the other hand, the effect of performance on reputation may also foster a scattering of forecasts if an occasional excellent forecast can give rise to a "superstar" effect. In both cases, the forecasts provided may deviate from the "best" forecast.[3]

1 The forecasters either work in the private sector or are professional economists working for universities and financial institutions in the respective country. Further information on the survey can be found on the website: www.consensuseconomics.com.

2 In addition, Consensus Economics also provides surveys of forecasts with a longer horizon on a less frequent basis. These are not taken into account here.

3 The question arises as to whether this anti-herding behavior is conditional on the publication of forecasters' affiliations. Pierdzioch and Rülke (2013) use the anonymous Livingston survey to show that interest rate forecasters deliberately place their forecasts away from the consensus forecast, i.e., anti-herd. Hence, anti-herding behavior does not appear to be conditional on the non-anonymity of a survey.

Finally, participation in each survey is voluntary, so that the composition of survey participants changes each time. If the survey terms required a certain continuity of participation, the forecasters could be forced to come up with new numbers even if they had not run a fully-fledged new forecast round recently.[4] A forecast that is close to the last consensus forecast may be on the safe side, thus providing incentives for herding behavior.

In order to examine the time series and cross-sectional dimensions of the survey data, Fig. 1 plots, for a selection of countries, the time series of (i) the cross-sectional mean values of the current-year GDP growth forecasts in each month (dashed lines), (ii) the actual annual GDP growth rates taken from the IMF database (solid lines), and (iii) the cross-sectional heterogeneity of forecasts, as measured by the cross-sectional range of forecasts (shaded areas). We show countries at various different stages of development, ranging from the United Kingdom and the United States to a selection of emerging countries that have previously been subject to severe economic crises, such as Argentina, Russia, and Turkey. The recession in the course of the global financial crisis is clearly visible in all cases. The vertical distance between the dashed line and the solid line can be interpreted as the cross-sectional forecast error.

These cross-sectional mean values move in tandem with the respective actual values, at least as far as the end-of-year values are concerned. This result is intuitive, because the forecast accuracy should increase as the forecast horizon shrinks. Another important observation is that the cross-sectional heterogeneity in forecasts is substantial. For instance, in January 1999, the forecast range of the growth rate in Russia was between −3.0% and −11.0%. In February 2009, the forecasts of the growth rate in the United States (United Kingdom) ranged between −3.1% and 2.1% (−3.5% and 2.5%), indicating a substantial degree of disagreement among the forecasters. Dovern, Fritsche, Loungani, and Tamirisa (2015) report a strong cross-sectional heterogeneity among business cycle forecasters. In addition, Dovern (2013) shows that the likelihood of forecast revisions depends on the forecast horizon, the business cycle, and strategic interactions between forecasters. Given this substantial heterogeneity, an analysis of the formation of GDP growth forecasts by individual professional forecasters seems appealing. As we will see, anti-herding behavior by growth forecasters is one source of the substantial heterogeneity in growth forecasts in most industrial countries.

3. Testing for forecaster herding

The intuition motivating the test for forecaster (anti-)herding developed by Bernhardt et al. (2006) can be explained best by considering as a benchmark scenario a forecaster who forms an "efficient" private forecast of the future GDP growth rate. The efficient private forecast should be unbiased and the probability of it overshooting or undershooting the growth rate should be 0.5, independently of the forecasts of others. In contrast, if forecasts are biased because forecasters (anti-)herd, the published forecasts will differ from the efficient private forecast. As a result, the probability of it overshooting or undershooting the growth rate should be larger or smaller than 0.5 (Pierdzioch et al., 2010).

In the case of forecaster herding, the published forecast deviates from the (unobservable) efficient private forecast because a forecaster follows the forecasts of others, measured in terms of the consensus forecast $\tilde{s}_{i,t-1,t+k}$ (the arithmetic mean forecast made by other forecasters).[5] Thus, the published forecast is closer to the consensus forecast than the efficient private forecast. A published forecast that exceeds the consensus forecast is smaller than the efficient private forecast, leaving less room for forecast overshooting. Similarly, a published forecast that is smaller than the consensus forecast is larger than the efficient private forecast, leaving less room for forecast undershooting. Conversely, in the case of forecaster anti-herding, a published forecast larger (smaller) than the consensus forecast is also larger (smaller) than the efficient private forecast, thus inflating the probability of overshooting (undershooting).

More formally, under the null hypothesis of unbiasedness (no herding or anti-herding), published forecasts should imply that the conditional probability of overshooting (undershooting), P_o (P_u), is 0.5, regardless of the consensus forecast. We then have:

$$\begin{aligned} P_o &= P(s_{t+k} < s_{i,t,t+k} \,|\, s_{i,t,t+k} > \tilde{s}_{i,t-1,t+k}, s_{t+k} \neq s_{i,t,t+k}) \\ &= 0.5, \end{aligned} \tag{1}$$

$$\begin{aligned} P_u &= P(s_{t+k} > s_{i,t,t+k} \,|\, s_{i,t,t+k} < \tilde{s}_{i,t-1,t+k}, s_{t+k} \neq s_{i,t,t+k}) \\ &= 0.5, \end{aligned} \tag{2}$$

where $s_{i,t,t+k}$ is the forecast made by forecaster i in period t for the growth rate s_{t+k} in period $t + k$ (with $k = 12, 11, \ldots, 1$ for the current-year forecasts, and $k = 24, 23, \ldots, 13$ for the next-year forecasts), and $\tilde{s}_{i,t-1,t+k}$ refers to the last observable consensus forecast. Forecaster herding then implies:

$$\begin{aligned} P_o &= P(s_{t+k} < s_{i,t,t+k} \,|\, s_{i,t,t+k} > \tilde{s}_{i,t-1,t+k}, s_{t+k} \neq s_{i,t,t+k}) \\ &< 0.5, \end{aligned} \tag{3}$$

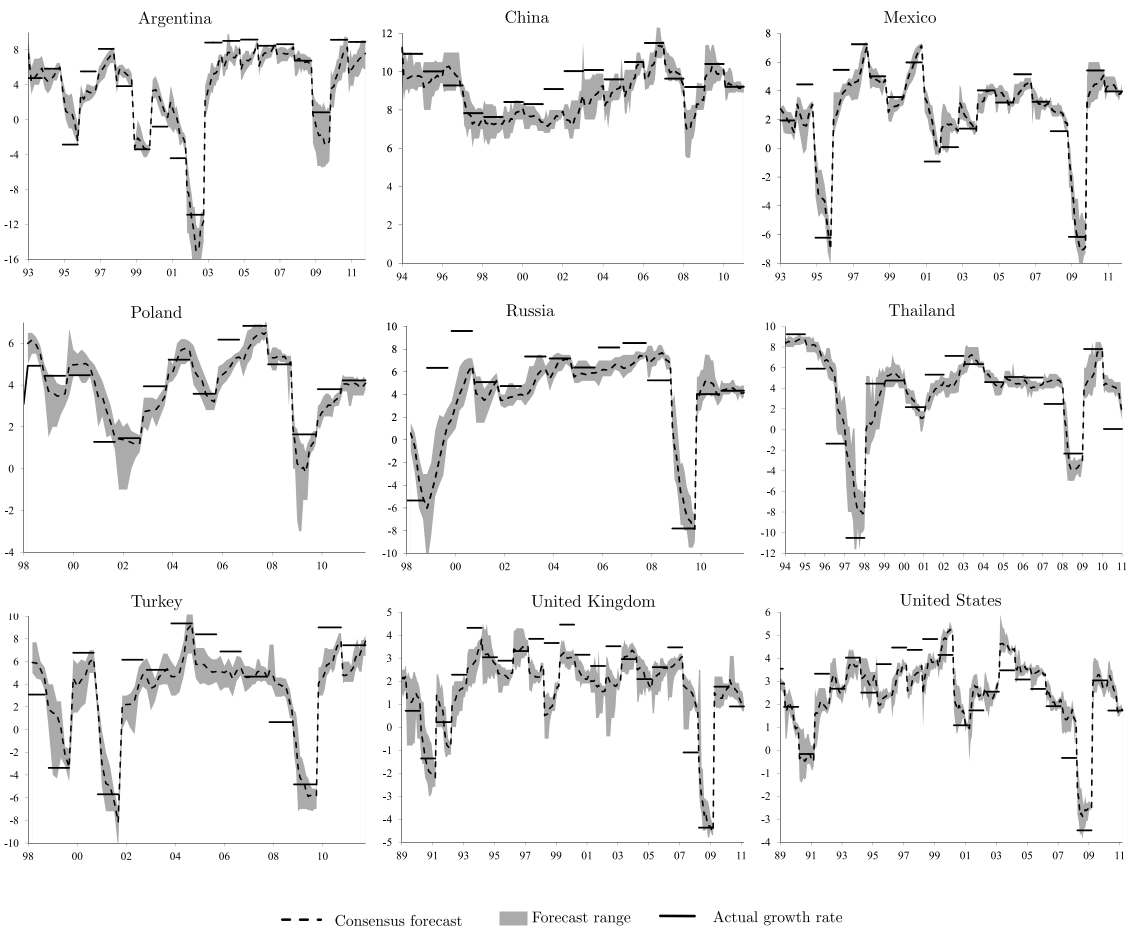
$$\begin{aligned} P_u &= P(s_{t+k} > s_{i,t,t+k} \,|\, s_{i,t,t+k} < \tilde{s}_{i,t-1,t+k}, s_{t+k} \neq s_{i,t,t+k}) \\ &< 0.5. \end{aligned} \tag{4}$$

In the case of forecaster anti-herding, P_o and P_u in Eqs. (3) and (4) exceed 0.5.

Bernhardt et al. (2006) suggest that forecaster (anti-)herding be tested for using the average of the two conditional probabilities. Their test statistic is defined as $S = (P_o + P_u)/2$, and has an asymptotically normal sampling distribution. Unbiased forecasts imply $S = 0.5$,

[4] A survey among participants in the ECB Survey of Professional Forecasters (SPF) (ECB, 2009) revealed that about 60% of the forecasters involved run full-fledged updates of their GDP growth forecasts only on a quarterly basis or even less frequently, but may provide monthly forecast updates to the SPF nevertheless.

[5] We used the consensus forecast from the previous survey so as to ensure that a forecaster's information set includes the last published consensus forecast. In addition, we also excluded the forecaster's own projection published in the previous survey, in order to compute a forecaster-specific consensus forecast.



Notes: The graphs show the cross-sectional consensus current-year forecast (dotted line), the forecast range (shaded area), and the actual values (solid line). The actual growth rates are taken from the database of the International Monetary Fund.

Fig. 1. Actual and expected GDP growth rates for selected countries.

herding implies $S < 0.5$, and anti-herding implies $S > 0.5$. To calculate the test statistic for our baseline result, we used all of the available surveys and pooled the forecasters for each country and among all forecasters.

The idea of applying the test proposed by Bernhardt et al. (2006) to our research question is appealing for a number of reasons. First, the test does not depend on any assumptions about the distributions of forecasts or forecast errors. This is simply because the test statistic does not depend on the degree of overshooting or undershooting, and hence, does not depend on the distribution of the forecasters' priors. Second, Bernhardt et al. (2006) show that the distribution of the realized values is not assumed to follow a normal distribution either. In fact, the test is conservative, in the sense that asymmetric shocks in the realized values only increase the variance of the test statistic, leaving the S statistic itself unaffected.[6] The test statistic, S, is also robust to both outliers in the data and data entry errors, because the probabilities P_u and P_o are computed as the relative frequencies of events from a large number of forecasts. Third, the test statistic is robust even when the forecasts are built on an asymmetric loss function.[7] Fourth, the test statistic S has an asymptotic normal distribution and is robust to various problems that may arise in the case of, e.g., correlated forecast errors, market-wide shocks, and optimism or pessimism among forecasters. It is the averaging that makes the test robust under the null hypothesis. Finally, it is relevant for our analysis to note that outliers and large disruptive events, such as the global financial crisis of 2008–09, should have only a minor effect on the conditional probabilities (i.e., empirical frequencies of events). The test statistic S is robust to such events.

The null hypothesis of the test is that – in the absence of herding or anti-herding behaviors – forecasters target

[6] The reason for this is that, for example, a market-wide shock that drives up the growth rate increases P_o and decreases P_u, leaving their average unaffected.

[7] Patton and Timmermann (2007) and Elliott, Komunjer, and Timmermann (2005) provide evidence that macroeconomic forecasts are based on asymmetric loss functions.

the actual value. Of course, the assumption that the private forecast is unbiased in the absence of herding or anti-herding behaviors could be challenged. One could argue that growth forecasters may have a positive forecast bias because they take into account the positive external effects of optimistic forecasts on financial products. Bernhardt et al. (2006) themselves raise the possibility that financial analysts may issue optimistic forecasts early in the forecasting cycle and pessimistic forecasts later in the cycle, a phenomenon that is unrelated to herding. Because of the specific design of the S statistic that nets out opposing effects, the test statistic is unaffected by this type of bias.

4. Empirical results and robustness tests

In this section, we present our baseline results and focus on various macroeconomic (economic crises and forecaster uncertainty) and microeconomic (number of competing forecasters, forecasters' affiliations, and forecasters' performances) factors that might determine whether or not forecasters show a tendency to (anti-)herd. Furthermore, we also present a set of robustness tests, such as results based on real-time data, a sample split analysis, and results based on unemployment forecasts.

4.1. Baseline results

Table 1 presents our main results and reports the S statistic, its standard errors, and the boundaries of a confidence band for all countries. It also contains information on the sample period, the numbers of forecasters available for every country in our sample, and the total numbers of observations. Three things stand out. First, the test statistic concerning the current-year (next-year) forecasts exceeds the value of 0.5 significantly in 20 (31) cases out of 45, indicating anti-herding behavior in these cases. Our results suggest that business cycle forecasters scatter their projections away from the consensus forecast more often as the forecast horizon lengthens. This finding supports the results of Lamont (2002), who reports that the disagreement among forecasters is due to strategic behavior. He emphasizes that the forecaster's age is related positively to the likelihood of bold forecasts. In our context, a series of bold forecasts reflects anti-herding behavior by forecasters.[8]

Second, for nine (five) countries, the current-year (next-year) test statistic is significantly lower than 0.5, implying that business cycle forecasters in some countries place their forecasts close to the consensus. Such a herding behavior is observed for Argentina, Colombia, Estonia, Latvia, Lithuania, Mexico, Peru, Poland, and Turkey. These are mostly developing or transition countries. Third, for some countries, such as Brazil, Italy, Sweden and the Ukraine, the S statistic is not significantly different from 0.5, indicating no (anti-)herding.

In summary, Table 1 gives the impression that forecaster herding is more likely to be observed in GDP growth forecasts for emerging market economies. A simple regression of the S statistic on GDP per capita (as of 2011 for all 45 countries) reveals a positive but statistically insignificant relationship between the two. Hence, it is likely that factors other than the stage of economic development of the respective economy are responsible for the observed herding behavior in GDP growth forecasts. Since less-developed countries are also more prone to economic and financial crises, one may therefore hypothesize that it is the incidence of economic turbulence that causes the differences in (anti-)herding behavior, rather than the level of economic development. We will test this hypothesis more formally in Section 4.2.

Given the large swings and sharp reversals in the growth rate during our sample period, we also analyzed the variation in the S statistic over time. Fluctuations in the S statistic should signal changes in the prevalence of forecaster (anti-)herding. To analyze the (anti-)herding behavior over time, we use a rolling-window estimation approach and pool the forecasts for every country within one year.[9] Fig. 2 plots the S statistics and the resulting 99% confidence bands for some selected countries. The figure also includes a crisis index developed by Reinhart and Rogoff (2011), as additional information.[10] The results suggest that the S statistic varies over time, sometimes significantly exceeding the value of 0.5 (anti-herding) and sometimes falling short of it (herding). Overall, anti-herding behavior is more prevalent than herding, especially during economically quiet times, indicating that forecasters deliberately differentiate their forecasts from those of their colleagues.[11]

Interestingly, the S statistic and the RR index seem to exhibit contradictory movements. For instance, during the Russian crisis in 1998–1999, the herding statistic is significantly lower than 0.5, indicating forecaster

[8] Rather than splitting the sample into current-year and next-year forecasts, we also tested explicitly whether the degree of (anti-)herding varies with the forecast horizon ($k = 24, 23, \ldots, 2, 1$). The results at a country level, which are available upon request, support our main finding of the anti-herding behavior of business cycle forecasters. Interestingly, for some countries (Australia, China, Japan, Taiwan, United Kingdom and United States), we observe a significant drop in the test statistic for four-quarter-ahead forecasts ($k = 12$). This might be related to a carry-over effect (Clements, 2014), i.e., the switch from next-year to current-year forecasts.

[9] This is a reasonable strategy for comparing cross-country differences but could potentially be problematic since forecasts with different forecast horizons are being aggregated. However, results based on a survey conducted in a specific calendar month, e.g., January of any year, are similar to our baseline results and are available upon request.

[10] This index reflects a comprehensive crisis definition, as it encompasses financial, currency and economic crises. A composite indicator is constructed by adding up the incidence of crises within one year in each of six different domains (banking crisis, currency crash, domestic default (or restructuring), external default (or restructuring), inflation crisis, and stock market crash). The index ranges between zero and six, with six indicating a more severe crisis.

[11] To account for the possibility that business cycle forecasters target the first data release of GDP rather than the revised data value (which is typically published with a considerable time lag), we also calculated these results based on the real-time database of the OECD. This dataset is limited to 24 countries and the time period January 1999–December 2011. The results support our baseline results: for most specifications (40 out of 48), business cycle forecasters show anti-herding behavior, i.e., they have a test statistic which is significantly larger than 0.5, indicating that forecasters differentiate their forecasts from the consensus forecast.

Table 1
Empirical results.

Country (Period)	Year	S-stat.	Std. error	Lower 99%	Upper 99%	Obs.	No.
Argentina	Current	0.485**	0.01	0.462	0.508	3,306	53
(1993–2011)	Next	0.465*	0.01	0.440	0.490	2,769	53
Australia	Current	0.545*	0.01	0.524	0.565	4,155	41
(1990–2011)	Next	0.661*	0.01	0.640	0.682	3,965	41
Brazil	Current	0.503	0.01	0.479	0.527	3,059	53
(1993–2011)	Next	0.515	0.01	0.490	0.540	2,829	53
Bulgaria	Current	0.660*	0.02	0.600	0.720	481	13
(1998–2011)	Next	0.582*	0.02	0.529	0.635	609	13
Canada	Current	0.495	0.01	0.474	0.516	4,010	37
(1999–2011)	Next	0.534*	0.01	0.513	0.556	3,812	37
Chile	Current	0.525*	0.01	0.501	0.550	2,936	43
(1993–2011)	Next	0.497	0.01	0.471	0.522	2,699	43
China	Current	0.543*	0.01	0.521	0.565	3,572	46
(1994–2011)	Next	0.601*	0.01	0.578	0.624	3,295	46
Colombia	Current	0.496	0.01	0.464	0.527	1,747	34
(1997–2011)	Next	0.474**	0.01	0.440	0.507	1,567	34
Croatia	Current	0.580*	0.02	0.521	0.639	533	12
(1998–2011)	Next	0.517	0.02	0.453	0.581	427	12
Czech Republic	Current	0.511	0.01	0.481	0.541	1,872	32
(1998–2011)	Next	0.553*	0.01	0.520	0.585	1,674	32
Estonia	Current	0.428*	0.02	0.368	0.487	492	13
(1998–2011)	Next	0.419*	0.03	0.351	0.486	393	13
France	Current	0.568*	0.01	0.548	0.587	4,741	40
(1999–2011)	Next	0.553*	0.01	0.532	0.573	4,247	40
Germany	Current	0.525*	0.01	0.510	0.541	7,197	54
(1999–2011)	Next	0.505	0.01	0.489	0.521	6,630	54
Hong Kong	Current	0.519**	0.01	0.496	0.541	3,284	44
(1994–2011)	Next	0.530*	0.01	0.506	0.554	3,006	44
Hungary	Current	0.487	0.01	0.455	0.519	1,725	29
(1998–2011)	Next	0.580*	0.01	0.546	0.614	1,560	29
India	Current	0.516	0.01	0.490	0.542	2,506	45
(1994–2011)	Next	0.587*	0.01	0.559	0.616	2,124	45
Indonesia	Current	0.583*	0.01	0.558	0.608	2,736	46
(1994–2011)	Next	0.591*	0.01	0.565	0.617	2,553	46
Italy	Current	0.498	0.01	0.476	0.520	3,625	40
(1999–2011)	Next	0.499	0.01	0.476	0.522	3,377	40
Japan	Current	0.598*	0.01	0.580	0.616	5,263	45
(1999–2011)	Next	0.607*	0.01	0.587	0.628	4,228	45
Latvia	Current	0.459**	0.02	0.398	0.519	474	13
(1994–2011)	Next	0.440**	0.03	0.372	0.507	383	13
Lithuania	Current	0.435*	0.02	0.371	0.498	434	12
(1994–2011)	Next	0.460**	0.03	0.389	0.531	347	12
Malaysia	Current	0.518**	0.01	0.494	0.541	3,122	52
(1994–2011)	Next	0.554*	0.01	0.529	0.578	2,919	52
Mexico	Current	0.460*	0.01	0.437	0.482	3,492	50
(1993–2011)	Next	0.543*	0.01	0.520	0.566	3,239	50
Netherlands	Current	0.498	0.01	0.467	0.529	1,782	29
(1995–2011)	Next	0.542*	0.01	0.510	0.575	1,641	29
New Zealand	Current	0.584*	0.01	0.559	0.610	2,668	27
(1994–2011)	Next	0.608*	0.01	0.581	0.634	2,485	27
Norway	Current	0.549*	0.01	0.515	0.583	1,489	20
(1998–2011)	Next	0.573*	0.01	0.538	0.609	1,400	20
Peru	Current	0.461*	0.01	0.429	0.493	1,726	33
(1997–2011)	Next	0.490	0.01	0.456	0.523	1,526	33
Philippines	Current	0.499	0.02	0.440	0.558	491	18
(2009–2011)	Next	0.595*	0.03	0.518	0.672	293	18
Poland	Current	0.471*	0.01	0.441	0.500	1,979	36
(1998–2011)	Next	0.497	0.01	0.465	0.528	1,783	36
Romania	Current	0.515	0.01	0.477	0.553	1,194	29
(1998–2011)	Next	0.528**	0.02	0.488	0.568	1,078	29
Russia	Current	0.474**	0.01	0.442	0.505	1,733	40
(1998–2011)	Next	0.565*	0.01	0.532	0.599	1,561	40
Singapore	Current	0.510	0.01	0.486	0.534	2,978	44
(1994–2011)	Next	0.536*	0.01	0.511	0.561	2,787	44

(continued on next page)

Table 1 (*continued*)

Country (Period)	Year	*S*-stat.	Std. error	Lower 99%	Upper 99%	Obs.	No.
Slovakia	Current	0.506	0.01	0.468	0.544	1,204	27
(1998–2011)	Next	0.583*	0.02	0.542	0.623	1,090	27
Slovenia	Current	0.453**	0.02	0.397	0.510	561	13
(1998–2011)	Next	0.538	0.02	0.474	0.602	451	13
South Korea	Current	0.578*	0.01	0.555	0.601	3,186	40
(1994–2011)	Next	0.524**	0.01	0.499	0.548	2,842	40
Spain	Current	0.544*	0.01	0.518	0.569	2,694	28
(1995–2011)	Next	0.546*	0.01	0.518	0.573	2,355	28
Sweden	Current	0.504	0.01	0.478	0.529	2,718	32
(1995–2011)	Next	0.490	0.01	0.464	0.516	2,533	32
Switzerland	Current	0.564*	0.01	0.534	0.593	2,016	19
(1998–2011)	Next	0.553*	0.01	0.522	0.584	1,842	19
Taiwan	Current	0.513	0.01	0.489	0.537	2,968	36
(1994–2011)	Next	0.566*	0.01	0.541	0.592	2,733	36
Thailand	Current	0.508	0.01	0.483	0.533	2,741	45
(1994–2011)	Next	0.537*	0.01	0.511	0.564	2,475	45
Turkey	Current	0.456*	0.01	0.424	0.488	1,664	38
(1998–2011)	Next	0.467**	0.01	0.431	0.502	1,419	38
Ukraine	Current	0.499	0.01	0.463	0.534	1,352	28
(1998–2011)	Next	0.505	0.01	0.466	0.544	1,163	28
United Kingdom	Current	0.543*	0.01	0.528	0.559	7,516	68
(1999–2011)	Next	0.599*	0.01	0.583	0.614	7,241	68
United States	Current	0.551*	0.01	0.535	0.567	6,941	65
(1999–2011)	Next	0.535*	0.01	0.518	0.551	6,334	65
Venezuela	Current	0.518**	0.01	0.492	0.544	2,551	42
(1993–2011)	Next	0.597*	0.01	0.569	0.624	2,253	42
Financial	Current	0.519*	0.002	0.515	0.524	85,583	1,127
forecasters	Next	0.543*	0.002	0.539	0.548	78,154	1,127
Non-financial	Current	0.509*	0.003	0.502	0.517	32,442	450
forecasters	Next	0.532*	0.003	0.525	0.540	28,970	450

Notes: For each country, the table reports the test statistic (*S*-stat.) based on pooling the data among all forecasters and all surveys. Std. error refers to the standard error. Obs. is the number of forecasts. No. reflects the numbers of forecasters who have submitted forecasts.
* indicates significance at the 1% confidence level ($H_0: S = 0.5$).
** indicates significance at the 10% confidence level ($H_0: S = 0.5$).

herding behavior. Similar evidence is found for Argentina (2000–2003), Mexico (1994–1996 and 1999–2000) and Turkey (1999–2001): whenever the RR crisis index increases, forecasters tend to demonstrate herding. This was also true to a lesser extent for industrial countries in the course of the global financial crisis, e.g., in the United Kingdom. Overall, Fig. 2 corroborates our hypothesis that there is a link between herding behavior and times of elevated economic uncertainty, rather than a link between herding behavior and the level of economic development. We will investigate this hypothesis more thoroughly in the next section.[12]

4.2. Macroeconomic determinants of forecaster (anti-)herding

In this section, we analyze the link between business cycle forecasters' (anti-)herding behaviors and the incidence of difficult economic times. To this end, we do not restrict ourselves to crisis periods but investigate times of high economic uncertainty more broadly. To ensure the robustness of our results, we define economic uncertainty from three completely different angles.

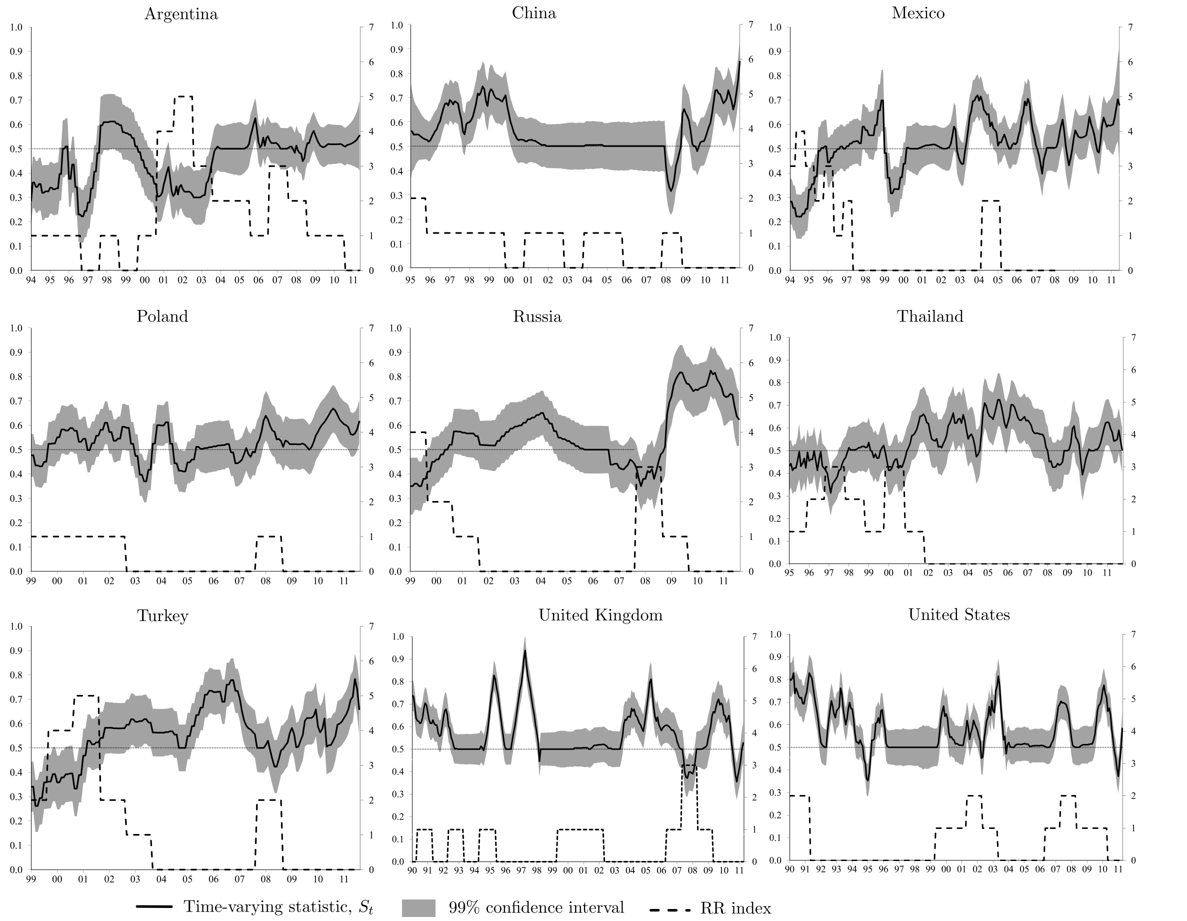
Our first measure is the RR index described above, which defines financial, currency and economic crises explicitly. It is a discrete variable ranging between zero and six, with a higher value indicating a more severe crisis, in the sense that it affects many dimensions of the economy.

Our second measure identifies recessions according to the conventional rule, namely two consecutive quarters of negative quarter-on-quarter GDP growth. Thus, this measure of uncertainty refers to recessions rather than economic crises. We construct a binary dummy for each year that takes the value of one if two consecutive quarters exhibit negative growth rates, and zero otherwise.[13]

As a third measure of economic uncertainty, we use a proxy that is observed from the Consensus Economics survey directly. To measure the extent of the macroeconomic uncertainty, we make use of the forecasting cycle within this survey, which consists of 24 months for each year. Since the forecasters are required to predict the growth rates of the current and the next year, each year is forecast

[12] As a further robustness check, we repeated our exercise focussing only on the time period of the so-called 'Great Moderation'. To this end, we only used data up to 2007. The results, which are available upon request, show that, for most countries, the test statistic significantly exceeds the unbiased value of 0.5. However, countries like Australia, Japan, and New Zealand still show a relatively high value of the test statistic, while countries like Argentina, Poland, and Turkey show test statistic results that indicate herding behavior. Hence, focusing on the Great Moderation does not seem to affect our results.

[13] If negative growth is recorded in the fourth quarter of a year and the first quarter of the subsequent year, we date the recession to the subsequent year.



Notes: The graphs show the unbiased value of 0.5 (fine dotted line), the Reinhart-Rogoff index (dotted line), the time-varying S_t statistic (solid line), and the 99% confidence interval (shaded area).

Fig. 2. Time-varying herding statistic and the RR index, S_t.

in 24 consecutive surveys. Naturally, the closer the survey approaches to the end of the forecasting cycle, the more precise the forecasts are. This can be seen in Fig. 1, which shows that the consensus forecast (dotted line) approaches the realized value (solid line) as far as end of year forecasts are concerned. The extent of the revisions for a specific year is then a natural measure of the level of forecast uncertainty. We proceed by using the sequence of the 24 consecutive consensus forecasts within a forecasting cycle and calculate the standard deviation for each forecaster over time as a proxy for forecast revisions. Hence, we obtain a measure of the forecast uncertainty for each year and country. In total, we have more than 7000 observations for each of the available measures. Unlike the other two measures, this proxy interprets economic uncertainty symmetrically, i.e., high forecast revisions in boom periods are treated in the same way as similar revisions during recessions. Furthermore, it is not a binary dummy but a continuous variable.

Table 2 reports the correlations between the S statistic and the three measures of economic uncertainty. The level of correlation between our three measures is significantly different from zero but quite low, which is not surprising, given that they are derived from totally different data sources and statistical approaches, and also define economic uncertainty from different points of view. Nevertheless, the correlation between the measures and the S statistic is negative in all three cases and significant in two of them, indicating that the S statistic declines with the level of economic uncertainty. This confirms our previous visual analysis.

To underpin our argument further, we subsequently link all three measures of the forecast uncertainty to the time-varying (anti-)herding statistic $S_{c,t}$ for each country c at time t (recall that the time-varying S_t is shown in Fig. 2 for selected countries). To this end, we regress the $S_{c,t}$ statistic on each measure of the forecast uncertainty by means of the following equation:

$$S_{c,t} = \alpha_c + \beta Crisis_{c,t} + \gamma GDPpc_{c,t} + \epsilon_{c,t}, \qquad (5)$$

where the α coefficient indicates the presence of (anti-)herding behavior and the β coefficient measures the extent to which individual forecasters deviate from this behavior in more turbulent periods. Table 3 reports the estimation results based on OLS, using Newey and West (1987) robust standard errors to take autocorrelation and cross-section

Table 2
Correlation of crisis variables.

	S-statistic	Reinhart-Rogoff	Recession dummy	Forecast revisions
S-statistic	1	–	–	–
Reinhart-Rogoff	−0.063*	1	–	–
Recession dummy	−0.019	0.209*	1	–
Forecast revisions	−0.066*	0.233*	0.494*	1

Notes: The table reports Bravais-Pearson correlation coefficients.
* indicates significance at the 1% level.

Table 3
Relationship between (anti-)herding and crises.

Estimator	Reinhart-Rogoff		Recession		Revisions		GDPpc		RR & GDPpc	
	OLS	FE	OLS	FE	OLS	FE	OLS	FE	OLS	FE
α	0.574**	0.575**	0.573**	0.573**	0.575**	0.575**	0.565**	0.565**	0.571**	0.570**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.01)	(0.01)
β	−0.074**	−0.096**	−0.048	−0.043	−0.054*	−0.052**	–	–	−0.071**	−0.096**
	(0.03)	(0.01)	(0.08)	(0.03)	(0.02)	(0.01)	(–)	(–)	(0.03)	(0.01)
γ	–	–	–	–	–	–	0.023	0.021	0.012	0.003
	(–)	(–)	(–)	(–)	(–)	(–)	(0.03)	(0.02)	(0.03)	(0.02)
Obs.	7403	7403	6464	6464	7317	7317	7403	7403	7403	7403

Notes: The table reports the estimation results of Eq. (5): $S_{c,t} = \alpha_c + \beta Crisis_{c,t} + \gamma GDPpc_{c,t} + \epsilon_{c,t}$, where c (t) is a country (time) index. The results are based on either the Newey and West (1987) panel estimator or the fixed-effects estimator (FE), and robust standard errors are given in parentheses.
* indicates significance at the 10% significance level.
** indicates significance at the 1% significance level.

correlation into account, as well as in a panel fixed-effect specification to allow for systematic differences among the countries. We generally observe anti-herding behavior in normal times, as is reflected by the constant term α of about 0.57, which is significantly higher than 0.5, indicating that, on average, business cycle forecasters anti-herd. The coefficient on the crisis dummy suggests that there is a statistically significant negative relationship between forecaster anti-herding and the extent of the forecast uncertainty. The same result emerges if we replace the RR index with the standard deviation of forecast revisions. Again, the β coefficient is significantly negative, indicating that business cycle forecasters tend to herd in times of forecast uncertainty and anti-herd in normal times. With regard to the recession dummy, the β coefficient is negative and of a similar magnitude to that in the forecast revision case, but insignificant. Compared to this, the relationship between the time-varying S statistic and the GDP per capita, as measured by the γ coefficient, is not significantly different from zero. Also, when using both the RR index and the GDP per capita as explanatory variables, the coefficient on the crisis variable is still significant, while that on GDP per capita is not.

Thus, while forecasters typically anti-herd in normal times, they show a tendency towards herding at times of higher economic uncertainty. The value of about −0.07 in the RR specification, for example, suggests that a crisis of level 2 according to the RR index would be enough for forecasters to switch to herding behavior. Concerning the forecast revisions, the coefficient of −0.05 indicates that if the standard deviation of the forecast revisions increases by about 0.1 over the forecasting cycle of 24 months, the herding statistic falls significantly short of the unbiased value of 0.5. This indicates that herding behavior is more likely, the more that forecasters have revised their forecasts.

Thus, our results show that herding and an elevated economic uncertainty are related; however, they do not allow us to engage in a more thorough analysis of the causality of this relationship. Based on our data, we cannot identify the underlying reasons for herding behavior in times of economic turbulence. A more comprehensive analysis could provide a further scrutiny of the role of model uncertainty in forecasting models and the effects of more frequent data revisions in uncertain times on forecasting behaviors. However, forecasting behaviors and economic uncertainty are clearly related, as our results show. Prati and Sbracia (2010) provide further evidence and show that a greater forecast uncertainty in the Consensus Economics dataset with respect to macroeconomic variables tends to be associated with currency crises.

4.3. Microeconomic determinants of forecaster (anti-)herding

4.3.1. Number of competing forecasters

At the micro-level, (anti-)herding behavior might also depend on the number of competing forecasters. Ottaviani and Sørensen (2006) argue that forecasters in a forecasting contest differentiate their predictions from those of their competitors. Hence, having a larger number of colleagues providing forecasts might motivate a forecaster to deliver more extreme forecasts, i.e., to anti-herd. In our data sample, the numbers of participants vary substantially across countries. While only seven forecasters participated in the survey for Lithuania in March 2011, 39 forecasters submitted growth forecasts for the July 1993 survey in the UK. To analyze whether the test statistic varies with the number of forecasters, we calculate it for each survey in each country and relate it to the number of participants in each survey. The results, which are available upon request, support our baseline result that, on average, the test statistic is above 0.5. The results also show that there is no clear-cut relationship between the number of participants and the test statistic. Hence, the number of competitors does not seem to affect (anti-)herding behavior.

4.3.2. Forecaster affiliation

In order to conduct a further analysis of microeconomic factors that may determine the (anti-)herding behavior, we take advantage of the fact that the survey includes the affiliations of the respective forecasters. Hence, we clustered the group of forecasters into those working with the financial sector (i.e., with investment banks, commercial banks, and consultancies), and those who are affiliated with non-financial companies (e.g., research industries, universities, and industry). As these two groups of forecasters respond to the needs of different customers, they may exhibit different strategic behaviors in their forecasts, as was suggested by the findings of Capistran and Timmermann (2009). Table 1 shows the test statistics for the two groups aggregated over all countries. Anti-herding behaviors are more pronounced for financial companies, but are not significantly different between the two groups, as the upper and lower bounds show. Interestingly, for both groups, anti-herding behavior is significantly more prevalent for next-year forecasts than for current-year forecasts. For individual financial (non-financial) forecasters, 455 (152) out of 1127 (450) forecasters, i.e., 40% (33%), have a test-statistic larger than 0.5, while only 266 (106), i.e., 24% (23 %), exhibit a test statistic which is significantly lower than 0.5. This indicates that forecasters tend to anti-herd even at the micro-level, but that there are no differences between financial and non-financial forecasters.

4.3.3. Forecasters' ex-ante and ex-post performances

(Anti-)herding behavior might also be related to ex-ante and ex-post forecast performances. By ex-ante forecast performance, we mean that a forecaster who has experienced large forecasting errors in the recent past may strategically bias his upcoming forecast towards the mean (i.e., show herding behavior) in order to re-gain credibility, or avoid worsening his reputation. In other words, we explain forecasting behaviors based on previous forecast performances at the micro level. By ex-post forecasting performance, we refer to the predictive accuracy of an individual forecaster's GDP growth forecasts conditional on his forecasting behavior (i.e., herding or anti-herding). Put differently, in this case we explain forecast performances based on strategic forecasting behaviors.

To analyze this relationship, we linked the (ex-ante) forecast performance to (anti-)herding behavior by regressing the test statistic for each survey and country on the forecast error (across all countries) of the previous survey. Interestingly, the tendency to herd increases with the previous forecast error. This indicates that forecasters tend to hide in the crowd once their forecast error is large, and supports our baseline result that forecasters prefer to herd in times of economic uncertainty.

A forecaster who differentiated his forecasts from the forecasts of his colleagues might have a different (ex-post) forecast accuracy. Laster et al. (1999) argue that forecasters may use extreme forecasts to attract the attention of customers, at the expense of the forecast accuracy. To examine the link between ex-post forecast performances and (anti-)herding, we calculated the individual test statistic and the root mean squared error for each and every forecaster as a measure of the forecast accuracy. The results indicate that there is only very weak evidence that anti-herding results in a lower forecast accuracy at the level of the individual forecaster.

5. Concluding remarks

This study investigates (anti-)herding behaviors of business cycle forecasters based on individual forecast data of participants in the Consensus Economics survey. The main finding is that forecasters show anti-herding behaviors in their predictions of GDP growth, at least in normal times and for industrialized countries. Thus, forecasters appear to make extreme forecasts in order to reap the "superstar bonus" if the forecast turns out to be the most accurate. However, the extent of anti-herding among the industrialized countries varies substantially. In contrast, for emerging economies, we find a tendency toward herding behaviors, as forecasters tend to move their forecasts towards the consensus.

The evidence of anti-herding behavior in industrial countries is particularly interesting given the recent empirical studies that have confirmed that it is difficult to beat the consensus forecast (Genre, Kenny, Meyler, & Timmermann, 2013). This would actually suggest herding behavior to be the optimal strategy for forecasters. The evidence of forecaster anti-herding indicates that the reputation effects of delivering extreme forecasts outweigh the incentives to optimize the forecast accuracy.

Furthermore, we find that the extent to which growth forecasters (anti-)herd varies over time. Based on a rolling window estimation, we find that forecaster (anti-)herding was somewhat less pronounced in the period from 2001 to 2003, and became more prevalent thereafter. Using three different indicators of economic uncertainty, a regression analysis confirms that forecasters show a tendency to approach the previous consensus forecast at more turbulent times. It seems that more uncertain economic circumstances induce professional forecasters to hide in the flock in order to avoid potential reputation losses arising from wrong individual forecasts. After all, forecast errors made by a majority of forecasters can be considered to be unavoidable, and therefore will not affect the reputations of individual forecasters. We conclude that the forecast heterogeneity is reduced artificially at times of high economic uncertainty, underrating the extent of the disagreement among forecasters. In contrast, survey forecasts under normal economic conditions report an overrated picture of forecast disagreement. Both policy makers and financial market participants should take this into account when attempting to infer the macroeconomic uncertainty from survey data.

Acknowledgments

We are grateful to Graham Elliott (the editor), Bruno S. Frey, Markus Knell, Maritta Paloviita, two anonymous referees, and participants at an OeNB internal seminar and the CIRET 2012 conference for very helpful comments and suggestions. We are grateful to Ingeborg Schuch for proofreading the manuscript. We thank the Euro Area Macroeconomic Division of the European Central Bank for providing the data set. This paper was written when Jan-Christoph Rülke was visiting researcher at the OeNB, and he gratefully acknowledges the hospitality of the OeNB.

References

Batchelor, R. (2001). How useful are the forecasts of intergovernmental agencies? The IMF and OECD versus the consensus. *Applied Economics*, *33*, 225–235.

Bernhardt, D., Campello, M., & Kutsoati, E. (2006). Who herds? *Journal of Financial Economics*, *80*, 657–675.

Bewley, R., & Fiebig, D. G. (2002). On the herding instinct of interest rate forecasters. *Empirical Economics*, *27*, 403–425.

Capistran, C., & Timmermann, A. (2009). Disagreement and biases in inflation expectations. *Journal of Money, Credit and Banking*, *41*(2–3), 365–396.

Clements, M. P. (2014). Forecast uncertainty—ex ante and ex post: US inflation and output growth. *Journal of Business and Economic Statistics*, *32*(2), 206–216.

Dovern, J. (2013). When are GDP forecasts updated? Evidence from a large international panel. *Economics Letters*, *120*, 521–524.

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

ECB (2009). Results of a special questionnaire for participants in the ECB Survey of Professional Forecasters.

Elliott, G., Komunjer, I., & Timmermann, A. G. (2005). Estimation and testing of forecast rationality under flexible loss. *Review of Economic Studies*, *72*, 1107–1125.

Freedman, A. J. (2013). *Mud on your face, big disgrace vs. the wisdom of crowds: macroforecasters' herding behavior over time*. Honors Projects, Paper 50.

Gallo, G. M., Granger, C. W. J., & Jeon, Y. (2002). Copycats and common swings: The impact of the use of forecasts in information sets. *IMF Staff Papers*, *49*(1), 4–21.

Genre, V., Kenny, G., Meyler, A., & Timmermann, A. (2013). Combining expert forecasts: Can anything beat the simple average? *International Journal of Forecasting*, *29*(1), 108–121.

Grisse, C. (2009). *Are expectations about economic activity self-fulfilling? An empirical test*. Federal Reserve Bank of New York, Working Paper.

Lamont, O. A. (2002). Macroeconomic forecasts and microeconomic forecasters. *Journal of Economic Behavior and Organization*, *48*(3), 265–280.

Laster, D., Bennett, P., & Geoum, I. S. (1999). Rational bias in macroeconomic forecasts. *Quarterly Journal of Economics*, *114*, 293–318.

Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica*, *55*, 703–708.

Ottaviani, M., & Sørensen, P. N. (2006). The strategy of professional forecasting. *Journal of Financial Economics*, *81*(2), 441–466.

Patton, A. J., & Timmermann, A. G. (2007). Testing forecast optimality under unknown loss. *Journal of the American Statistical Association*, *102*, 1172–1184.

Pierdzioch, C., & Rülke, J. C. (2013). A note on the anti-herding instinct of interest-rate forecasters. *Empirical Economics*, *45*(2), 665–673.

Pierdzioch, C., Rülke, J. C., & Stadtmann, G. (2010). New evidence of anti-herding of oil-price forecasters. *Energy Economics*, *32*(6), 1456–1459.

Pons-Novell, J. (2003). Strategic bias, herding behaviour and economic forecasts. *Journal of Forecasting*, *22*(1), 67–77.

Prati, A., & Sbracia, M. (2010). Uncertainty and currency crises: evidence from survey data. *Journal of Monetary Economics*, *57*, 668–681.

Reinhart, C. M., & Rogoff, K. S. (2011). From financial crash to debt crisis. *American Economic Review*, *101*(5), 1676–1706.

Scharfstein, D., & Stein, J. (1990). Herd behavior and investment. *American Economic Review*, *80*, 465–479.

Zarnowitz, V. (1984). The accuracy of individual and group forecasts from business outlook surveys. *Journal of Forecasting*, *3*(1), 11–26.

Zarnowitz, V., & Lambros, L. A. (1987). Consensus and uncertainty in economic prediction. *Journal of Political Economy*, *95*, 591–621.

Jan-Christoph Rülke is Professor for Economics at the EBS University. He has published in the *Journal of Banking and Finance, Journal of Economic Dynamics and Control, Journal of International Money and Finance*, and *Journal of Macroeconomics*, among others.

Maria Silgoner is Senior Economist at the Oesterreichische Nationalbank (OeNB). She has published in the *Journal of Common Market Studies, Applied Economics, Journal of Business Cycle Measurement* and *Analysis and China Economic Review*.

Julia Wörz is Senior Economist at the OeNB. She has published in the *Journal of International Economics, Economics of Transition, Review of World Economics, Journal of Industry, Competition and trade* and *Applied Economics*.