Received: 9 August 2020 | Revised: 28 October 2020 | Accepted: 2 December 2020

DOI: 10.1002/for.2764

RESEARCH ARTICLE

WILEY

Strategic bias and popularity effect in the prediction of economic surprises

Luiz Félix[1,2] | **Roman Kräussl**[3,4] | **Philip Stork**[1,5]

[1]Department of Finance, Vrije Universiteit Amsterdam, Amsterdam, Netherlands

[2]Multi-asset, APG Asset Management, Amsterdam, Netherlands

[3]Department of Finance, University of Luxembourg, Luxembourg, Luxembourg

[4]Hoover Institution, Stanford University, Stanford, California, USA

[5]Research field: Finance, Tinbergen Institute Amsterdam, Amsterdam, Netherlands

Correspondence
Philip Stork, Department of Finance, Vrije Universiteit Amsterdam, de Boelelaan 1105, 1081 HV Amsterdam, The Netherlands.
Email: p.a.stork@vu.nl

Abstract

Professional forecasters of economic data are remunerated based on accuracy and positive publicity generated for their firms. This remuneration structure incentivizes them to stick to the median forecast but also to make bold forecasts when they perceive to have superior private information. We find that skewness in the distribution of expectations, potentially created by bold forecasts, predicts economic surprises across a wide range of US economic indicators in-sample and out-of-sample, confirming our hypothesis that forecasters behave strategically and possess private information. This strategic bias found in US economic forecasts is also exhibited in individual forecasters' data as well as in continental Europe, the United Kingdom, and Japan. We show that it has been increasing both through time and in relation to the behavioral anchor bias. Our results suggest that the pervasiveness of the biases depends on the popularity of the economic indicator being released, both in the United States and internationally.

KEYWORDS

economic surprises, forecast error, predictability, skewness, strategic bias

JEL CLASSIFICATION

G14; F47; E44

1 | INTRODUCTION

> "...hey look, I make bold forecasts and specially in things like QE and massive events, yes, I am going to miss some things but I have the guts to make these bold calls and most people couch their forecasts, which, to me, makes them useless as economists...this is a golden opportunity, the best I have seen probably in my life-time ..."
>
> Economist Harry Dent live on Futures Now program at CNBC television,
>
> 8 December 2016

The median forecast of economic data is an important anchor of expectations for policy makers, governments, firms, and financial market participants. As expectations about the economy are a key input to policymakers' and private economic decision making, the consensus (i.e., the median forecast) plays a crucial role in the level of interest rates set by central banks, as well as the level of government spending, private investments, and financial market prices. The consensus is estimated via the aggregation of individual professional forecasters' economic predictions collected via surveys. Hence, it is one of the few economic measures that is ex ante in nature, rather than backward-looking information, which explains its importance.

Since professional forecasters may be subject to systematic biases, it is essential for users of consensus estimates to understand the factors influencing economic predictions. Biases in analysts' forecasts in a wider context is a widely investigated topic. The seminal research on the subject focuses on biases incurred by forecasters of earnings per share (FEPS) rather than in economic data. This literature mostly points to behavioral explanations for the biases.[1,2]

The literature on inefficiencies in the forecast of economic variables acknowledges strategic and behavioral reasons for the presence of systematic biases of forecasters. For instance, Scharfstein and Stein (1990) propose a 'reputational herding model' which suggests that forecasters (investment managers in their case) mimic the decision of others and ignore substantive private information, mostly due to concerns about their reputation in the labor market, which would cause forecasts to be concentrated around the consensus. Somewhat differently, Ottaviani and Sorensen (2006) investigate strategies carried out by professional forecasters, which lead to either forecasts that are excessively dispersed or forecasts that are biased towards the prior mean (herding), also due to reputational concerns. They conclude that in a 'winner-take-all contest', which is suggested to be the case for economic forecasting,[3] equilibrium forecasts are excessively differentiated. This equilibrium occurs because, even if reputational concerns cause agents to herd, when agents have substantial prior knowledge on their own superior forecasting ability, they tend to overweight the use of private information in their forecasting.

In the same line, Laster et al. (1999) develop a model in which forecasters have a dual goal: accuracy and publicity following from the wage schemes paid by employers to forecasters. Forecasters would signal confidence in their own forecasts by departing from the median and making them off-consensus when incentives related to their firms' positive publicity outpace the wages received by being as accurate as their peers. Note that posting a forecast close to the current median is, from a relative performance perspective, a safe course of action as the forecaster accuracy versus peers cannot, by concept, worsen. An important assumption of this 'signaling hypothesis' of Laster et al. (1999) is that firms value publicity given to the single top forecaster, as in a 'winner-take-all contest' assumed by Ottaviani and Sorensen (2006). In such context, forecasters also tend to excessively differentiate their estimates.

In the above literature, professional forecasters behave strategically to maximize wages or preserve their reputation, which are examples of strategic biases. Note that strategic biases are able to back both excessive dispersion and excessive concentration of forecasts.

Campbell and Sharpe (2009) is an important milestone on addressing economic forecasts from a behavioral bias perspective. Their study hypothesizes that the median forecast of individual economic releases is systematically biased towards the previous release. They argue that this bias is consistent with the adjustment heuristic, commonly known as anchoring, proposed by Tversky and Kahneman (1974). This cognitive bias is characterized by the human propensity to rely too heavily on the initial value (the 'anchor') of an estimation when updating forecasts.[4] Note, however, that the empirical consequence of anchoring (i.e., concentration of forecasts) is similar to the 'reputational herding model' suggested by Scharfstein and Stein (1990), which is a strategic bias rather than a cognitive one.

A drawback of the literature on economic forecasts is that their empirical conclusions cannot be generalized as they do not cover many indicators and geographies for the presence of the biases. For instance, the empirical test of Laster et al. (1999) is based only on the real gross domestic product/gross national product (GNP/GDP) forecasts for the United States, whereas Ehrbeck and Waldmann (1996) only consider the annualized discount rate on new issues of 91-day US Treasury bills. The analysis of Ashiya (2009) uses data from Japanese forecasters, who tend to continue herding rather than posting differentiated forecasts as they become older and more established (see Ashiya & Doi, 2001). However, the findings of Lamont (2002) are the opposite (i.e., dispersion of forecast rather than concentration) from Ashiya (2009) using US data, which is an important support for the 'signaling hypothesis' of Laster et al. (1999). Ashiya (2006) uses a wider cross-country data set for testing if economic revisions are rational, but his analysis relies on GDP and its deflator only. Campbell and Sharpe (2009) are the first to address bias in economic forecast using a more comprehensive set of economic indicators. They find that the previous economic releases of 10 important economic indicators explain up to 25% of the subsequent economic surprises. Their study is, however, based on US data only.

In this paper, we investigate whether the skewness in the distribution of forecasts of 43 economic indicators in the United States and 219 indicators internationally is linked to economic surprises. The main hypothesis of our paper is that the skewness in the distribution of forecasts contains information, and, hence, it is able to forecast economic surprises.

Finally, Legerstee and Franses (2015) use the number of forecasts collected as a predictor of future economic releases as a proxy for 'attention'. Arguably this explicit measure of popularity, which relates one to one to the number of forecasters, can be used as a direct predictor of economic data, as these authors do. Nevertheless, the number of forecasters can also be employed as a weighting scheme to test whether the pervasiveness of biases fluctuates with popularity, which is the approach we follow in this paper.

Our contribution to the literature on forecasting bias is threefold. First, we empirically establish that skewness in the distribution of forecasts present in a wide and global data set of economic expectations is able to significantly predict economic surprises. More specifically, we are the first to empirically validate the effects of the 'signaling hypothesis' of Laster et al. (1999) in a large multi-country data set of economic releases. In this hypothesis, forecasters behave strategically by making off-consensus forecasts as they possess superior private information (and they know a priori about their ability), which gets unveiled via the skewness of the distribution of forecasts.

Second, we provide evidence that the importance of the skewness in the distribution of forecasts in predicting economic surprises increases steadily through time and versus the anchor bias, the behavioral bias in economic forecasting proposed by Campbell and Sharpe (2009).

Third, by expanding the number of countries and indicators tested vis-à-vis the earlier literature on forecast biases and by using our popularity measure per economic indicator, we show that the prevalence of biases is related to the number of forecasters posting estimates per indicator. The same effect is observed when we compare our results for the United States to those in other countries, in which economic indicators are forecasted by much fewer analysts.

There are three key implications of our findings. First, they enable a better understanding of the informational content of the skewness in the distribution of economic forecasts by regulators, policy makers, and market participants. Second, our strong results provide evidence that the anchor bias is not the single bias widely found in economic forecasting. As a result, benchmarks for assessment of economic surprise models should be complemented with our suggested skewness measure. Third, the popularity effect identified supports the usage of a weighted scheme versus an unweighted one in the construction of economic surprise indexes. The popularity effect also reinforces the attention of economist and market participants to popular economic indicators.

The remainder of this paper is organized as follows. Section 2 describes the data and methodology. Section 3 presents our main empirical analysis. Section 4 checks for the robustness of our findings, and Section 5 concludes.

2 | DATA AND METHODOLOGY

We use economic release data from the ECO function in Bloomberg. These data comprise time-stamped real-time released figures for 43 distinct US economic indicators, as well as information on forecasters' expectations for each release; see Table 1. This expectations information comprises (1) the expected surprise conditional to the anchor, the ESA_t factor of Campbell and Sharpe (2009), which is based on previous economic release[5,6]; (2) the cross-sectional standard deviation of forecasts; (3) the lagged median survey expectations; and (4) the skewness in economists' forecasts, calculated as the mean minus the median survey expectations.

The anchor-based model of Campbell and Sharpe (2009) is specified by Equation 1, and it is used as the main benchmark for our experiments:

$$S_t = \alpha + \gamma ESA_t + \in_t, \tag{1}$$

The cross-sectional standard deviation of forecasts is employed as a measure of dispersion of forecasts, which is a control variable in our experiments based on the suggestion by Mankiw et al. (2003), Zhang (2006), and Capistran and Timmermann (2009) that the second moment of forecasts also reveals biases. The lagged median survey expectation is an additional control variable.

We use similar data sets for continental Europe, the United Kingdom, and Japan for robustness testing. Our daily data set spans the period from January 1997 to December 2016, covering 4,422 business days and 21,048 individual announcements.

We note that the economic indicators are released in different frequencies and throughout the month. This asynchronicity among indicators poses some challenges to process the information flow coming from them and to jointly test for the predictability of surprises. Therefore, predictability is separately tested for each indicator, and results are subsequently aggregated.

As we intend to use the state of the economy as a control variable, we implement the principal component analysis (PCA)[7]-based *nowcasting* method of Beber et al. (2015) using the same 43 distinct US economic indicators. Their *nowcasting* method allows us to access the real-time growth and inflation conditions present at

TABLE 1 Overview of US macro releases

#	Indicator name	Type	Start	Frequency	Release time	Direction	Stationary
1	US Initial Jobless Claims SA	Growth	31/12/96	W	14:30:00 GMT	−1	No
2	US Employees on Nonfarm Payroll	Growth	02/01/97	M	14:30:00 GMT	1	No
3	U-3 US Unemployment Rate Total	Growth	07/01/97	M	14:30:00 GMT	−1	No
4	US Employees on Nonfarm Payroll Manuf.	Growth	08/01/97	M	14:30:00 GMT	1	Yes
5	US Continuing Jobless Claims SA	Growth	09/01/97	W	14:30:00 GMT	−1	No
6	ADP National Employment Report	Growth	09/01/97	M	14:15:00 GMT	1	No
7	US Average Weekly Hours All Employees	Growth	10/01/97	M	14:30:00 GMT	1	No
8	US Personal Income MoM SA	Growth	10/01/97	M	14:30:00 GMT	1	Yes
9	ISM Manufacturing PMI SA	Growth	14/01/97	M	16:00:00 GMT	1	Yes
10	US Manufacturers New Orders Total	Growth	14/01/97	M	16:00:00 GMT	1	Yes
11	Federal Reserve Consumer Credit	Growth	16/01/97	M	21:00:00 GMT	1	No
12	Merchant Wholesalers Inventories	Growth	17/01/97	M	16:00:00 GMT	1	Yes
13	US Industrial Production MOM SA	Growth	17/01/97	M	15:15:00 GMT	1	Yes
14	GDP US Chained 2009 Dollars QoQ	Growth	28/01/97	Q	14:30:00 GMT	1	Yes
15	US Capacity Utilization % of Total	Growth	03/02/97	M	15:15:00 GMT	1	Yes
16	US Personal Consumption Expenditures	Growth	03/02/97	M	14:30:00 GMT	1	Yes
17	US Durable Goods New Orders Ind.	Growth	25/02/97	M	14:30:00 GMT	1	Yes
18	US Auto Sales Domestic Vehicle	Growth	04/03/97	M	22:59:00 GMT	1	No
19	Adjusted Retail & Food Service	Growth	26/03/97	M	14:30:00 GMT	1	Yes
20	Adjusted Retail Sales Less Autos	Growth	03/07/97	M	14:30:00 GMT	1	Yes
21	US Durable Goods New Orders Total	Growth	16/07/97	M	14:30:00 GMT	1	Yes
22	GDP US Personal Consumption Change	Growth	12/08/97	Q	14:30:00 GMT	1	Yes
23	ISM Non-Manufacturing PMI	Growth	26/11/97	M	16:00:00 GMT	1	No
24	US Manufacturing & Trade Inventories	Growth	12/12/97	M	16:00:00 GMT	−1	Yes
25	Philadelphia Fed Business Outlook	Growth	13/08/98	M	16:00:00 GMT	1	Yes
26	MNI Chicago Business Barometer	Growth	08/01/99	M	16:00:00 GMT	1	Yes
27	Conference Board US Leading Ind.	Growth	14/05/99	M	16:00:00 GMT	1	Yes
28	Conference Board Consumer Conf.	Growth	01/07/99	M	16:00:00 GMT	1	No
29	US Empire State Manufacturing	Growth	13/06/01	M	14:30:00 GMT	1	Yes
30	Richmond Federal Reserve Manuf.	Growth	13/06/01	M	16:00:00 GMT	1	Yes
31	ISM Milwaukee Purchasers Manuf.	Growth	28/12/01	M	16:00:00 GMT	1	Yes
32	University of Michigan Consumer Sent.	Growth	25/07/02	M	16:00:00 GMT	1	No
33	Dallas Fed Manufacturing Outlook	Growth	15/11/02	M	16:30:00 GMT	1	Yes
34	US PPI Finished Goods Less Food & En.	Inflation	30/01/03	M	14:30:00 GMT	1	Yes
35	US CPI Urban Consumers MoM SA	Inflation	30/04/04	M	14:30:00 GMT	1	Yes
36	US CPI Urban Consumers Less Food & En.	Inflation	26/05/05	M	14:30:00 GMT	1	Yes
37	Bureau of Labor Statistics Employment	Inflation	30/06/05	Q	14:30:00 GMT	1	Yes
38	US Output Per Hour Nonfarm Business	Inflation	25/10/05	Q	14:30:00 GMT	−1	Yes
39	US PPI Finished Goods SA MoM%	Inflation	02/08/06	M	14:30:00 GMT	1	Yes
40	US Import Price Index by End User	Inflation	31/07/07	M	14:30:00 GMT	1	Yes
41	US GDP Price Index QoQ SAAR	Inflation	05/02/08	Q	14:30:00 GMT	1	Yes
42	US Personal Con. Exp. Core MOM SA	Inflation	26/01/09	M	14:30:00 GMT	1	Yes

TABLE 1 (Continued)

#	Indicator name	Type	Start	Frequency	Release time	Direction	Stationary
43	US Personal Cons. Exp. Price YOY SA	Inflation	05/02/10	M	14:30:00 GMT	1	Yes

Note: This table reports the 43 US economic indicators used in our main analysis. Indicators are classified as either growth or inflation related. Column *Start* reports the date that the time series of each economic indicator begins. Column *Frequency* reports in which frequency the indicator is released, where *Q* stands for quarterly, *M* for monthly, and *W* for weekly. *Release time* reports the typical (most frequent) release time of the indicator in GMT time. *Direction* states the potential directional adjustment, represented by −1 when the given indicator reports a quantity that is inversely related to growth or inflation. The column *Stationary* shows if an indicator's series is stationary; a stationary adjustment (i.e., towards 6 months differences) is applied within our data manipulation step so the series can be modeled using our methodology.

the time of any economic release.[8] Table 1 provides details on stationarity adjustments, directional adjustments, frequency of release, starting publication date for the series, and (common) release time. Finally, we also use the 12-month change in stock market prices (i.e., the S&P500 index prices) and the VIX index to proxy for wealth effects and risk appetite, respectively, as additional control variables in our empirical analysis.

2.1 | Economic surprise predictive models

Following Equation 1, we extend the *anchor-only* predictive model for economic surprises by incorporating the skewness of the distribution of economic forecasts as well as other moments as control variables. The moments of the distribution of economic forecasts added are (1) the lagged median forecast (first moment); (2) the disagreement among forecasters (second moment); and (3) the skewness of forecasts (third moment). Equation 2 is our *unrestricted economic surprise* model (*UnES* model):

$$S_t = \alpha + ESA_\varphi + SurvLag_\varphi + Std_\varphi + Skewness_\varphi + Infl_\varphi + Growth_\varphi + Stocks_\varphi + VIX_\varphi + \in_t, \quad (2)$$

where subscript ϕ (used hereafter) is $t-1$, *ESA* is the expected surprise given anchor,[9] *SurvLag* is the lagged median forecast, *Std* is the dispersion (standard deviation) of economic estimates across forecasters, and *Skewness* is the skewness of economic estimates across forecasters. *SurvLag*, *Std* and *Skewness* are the three variables selected to test our hypothesis that alternative measures intrinsic to the pool of economic forecasts can reflect biases in expectations over economic releases. More specifically, we use *SurvLag* to test whether an anchor towards the previous median forecast exists. We employ *Std* to test for the effect of forecasters' disagreement and information uncertainty over the predictability of economic surprises, in line with Zhang (2006). *Skewness* is used to test for the presence of strategic behavior in economic forecasting, in line with the forecasters' dual-goal hypothesis of forecasting accuracy and publicity as discussed in Laster et al. (1999) and Ottaviani and Sorensen (2006). *Infl* and *Growth* are the states of inflation and economic growth produced by the *nowcasting* method implemented. *Stocks* and *VIX* are the stock market returns and implied volatility. *Infl*, *Growth*, *Stocks*, and *VIX* are control variables in our model.

3 | EMPIRICAL ANALYSIS AND RESULTS

We split our empirical analysis and results section into three parts. Section 3.1 reports the results in predicting economic surprises. Section 3.2 explores the presence of a popularity effect in the predictions. Section 3.3 tests our findings out-of-sample.

3.1 | Predicting economic surprises

In this section we report our findings from Equations 1 and 2, that is, the *anchor-only* (restricted) model and the *unrestricted* model, respectively, which we use to forecast economic surprises.

Table 2 reports aggregated results of these models across all 43 US economic indicators. We evaluate the sign consistency (with our expectations) and the statistical strength of the individual regressors by computing the percentage of times that the coefficients are positive (as expected) and statistically significant at the 10% level across regressions run separately for each economic indicator. The model quality is evaluated using explanatory power (R^2) as well as the Akaike information criteria (AIC) per individual (economic indicator's) regression.

The *anchor-only* model estimates confirm the general finding of the previous literature, in which the expected surprise given the anchor (*ESA*) is a strong predictor of economic surprises. We observe that *ESA* is significantly linked to surprises 65% of the times in our sample. This result is confirmed by the *unrestricted* model, in which *ESA* is statistically significant 67% of the times. The

TABLE 2 Aggregated results of anchor-only (restricted) and unrestricted economic surprise models for the United States

Region	United States	
Model	**Anchor-only**	**Unrestricted**
Panel A—Percentage of statistical significance per factor		
Intercept	0.35	0.42
Bias	0.65	0.67
Std		0.35
SurvLag		0.40
Skewness		0.72
Infl		0.16
Growth		0.33
Stocks		0.07
VIX		0.23
Panel B—Percentage of positive coefficients		
Intercept	0.47	0.56
Bias	0.81	0.88
Std		0.56
SurvLag		0.56
Skewness		0.93
Infl		0.49
Growth		0.30
Stocks		0.51
VIX		0.23
Panel C—Model quality		
Mean R^2	4%	17%
Median R^2	2%	14%
Stdev R^2	4%	10%
AIC	923	896

Note: Panel A reports the percentage of statistically significant coefficients across anchor-only and unrestricted regression models for economic surprises of US economic indicators. For example, 0.65 found for the *ESA* variable within the anchor-only model means that 65% of the *ESA* across the individual regressions run for the 43 US economic indicators are statistically significant at the 10% level. Panel B reports the percentage of positive coefficients across anchor-only and unrestricted regression models for economic surprises of US economic indicators. Panel C reports the mean, median, and standard deviation of the explanatory power (R^2) achieved across all indicator-specific regressions, as well as average Akaike information criteria (AIC).

results for the *unrestricted* model reveal that the *Skewness* factor is also often significant (72%) across our individual indicator regressions.

This finding supports our conjecture that forecasters may behave strategically by signaling confidence in their own forecasts (a strategic bias), which is in line with Laster et al. (1999).[10] *SurvLag* and *Std* are somewhat statistically significant, with 40% and 35% of the times, respectively. The result for *SurvLag* challenges our hypothesis that an anchor towards the previous median forecast holds empirically. The weak statistical significance of *Std* among our individual regressions also suggests that disagreement among forecasters and information uncertainty is linked to economic surprises. The control variables *Infl*, *Growth*, *Stocks*, and *VIX* are significant between 7% and 33% of times, suggesting a weak relation with economic surprises.

From an explanatory power perspective, the *unrestricted* model dominates the *anchor-only* model. The mean R^2 across the predictive surprise models of the different economic indicators is 4% for the *anchor-only* model and 17% for the *unrestricted* model (R^2 medians are 2% and 14%, respectively).

We report for the *anchor-only* regressions positive coefficients for the *ESA* factor 81% of the times. The *unrestricted* model delivers a positively signed *ESA* coefficient 88% of the times. Both results suggest a robust relationship between economic surprises and the anchor factor. The frequency of positive coefficients found for *Skewness* is, however, even higher than for *ESA*. The *Skewness* regressors are positive 93% of times across all regressions. *SurvLag* and *Std* are with 56% also largely positive but to a lesser extent than *Bias* and *Skewness*. Our control variables are to an even lesser extent positive (between 23% and 51%). The results provided by AIC are in line with R^2 as the average AIC for the *anchor-only* model is higher (926) than for the *unrestricted* model (896). These findings thus support our hypothesis that a strategic bias is embedded in economic forecasting due to strategic behavior of forecasters; a notion that is in line with Laster et al. (1999) and Ottaviani and Sorensen (2006).

Table 3 presents the results of the individual predictive surprise models (restricted and unrestricted). The R^2 *gain* ratio (reported in the last column) computes the number of times that R^2 of the *unrestricted* model exceeds the R^2 for the restricted model. From a R^2 perspective, the *unrestricted* models largely outperform the *anchor-only* model. The R^2 *gain* ratio ranges from 1 to infinity, as the average R^2 across the *anchor-only* model is 3.7%, whereas for the *unrestricted* model it is 17%.

The *Conference Board Consumer Confidence* indicator is the variable for which R^2 is the highest in the *anchor-only* model (14%), followed by the *US PPI Finished Goods SA Mom%* indicator (13%). Most R^2s are of a single digit level, and for only four indicators does the explanatory power exceed 10%. Most anchor coefficients are statistically significant at least at the 10% level.

When the *unrestricted* model is used, *US Personal Income MoM SA* (45%) is the indicator with the highest R^2, followed by *US GDP Price Index QoQ SAAR* (41%), and *Adjusted Retail & Food Service Sales* (36%). Most R^2s

TABLE 3 Results of anchor-only (restricted) and unrestricted models for economic surprises per economic indicator

Model	Anchor-only model				Unrestricted model				
Statistics/regressors	R^2	AIC	Intercept	Anchor	R^2	AIC	Intercept	Anchor	Std
US Initial Jobless Claims SA	0%	22,901	128.8	−0.1**	7%	21,925	−3586.0	−0.1	−0.2
US Employees on Nonfarm Payrol	0%	6151	−12**	−0.1	8%	5884	50**	0.0000	−0.001**
U-3 US Unemployment Rate Total	0%	−2451	0.0***	0.1	11%	−2362	0.0	0.0	0.1
US Employees on Nonfarm Payrol	2%	5003	−5009***	−0.2*	16%	4937	1235	0	−1***
US Continuing Jobless Claims S	2%	17,944	2	0.3***	8%	17,910	30***	0***	0
ADP National Employment Report	1%	3132	3786.5	0.1	13%	3129	49**	0	0
US Average Weekly Hours All Em	0%	−183	0.0	−0.1	27%	−196	0.0	0.4*	0.8*
US Personal Income MoM SA	3%	−2153	0.0**	0.1**	45%	−2109	0.0**	0.2***	0.2
ISM Manufacturing PMI SA	1%	996	0.1	0.2*	5%	939	2.2	0.1	0.1
US Manufacturers New Orders To	1%	−1777	0.0	0.0	7%	−1638	0.0*	0.1*	−0.4**
Federal Reserve Consumer Credi	1%	11,472	668.9*	0.1	8%	10,667	964	0	0
Merchant Wholesalers Inventori	1%	−1908	0.0**	0.1*	9%	−1796	0.0	0.2**	−0.2
US Industrial Production MOM S	7%	−2056	0.0*	0.2***	24%	−1940	0.0	0.3***	−0.4
GDP US Chained 2009 Dollars Qo	1%	−1860	0.0	0.0*	9%	−1787	0.0**	0.1***	0.3
US Capacity Utilization % of T	3%	−2063	0.0	0.2**	14%	−1928	0.0	0.3***	0.5*
US Personal Consumption Expend	9%	−2414	0.0	0.1***	15%	−2248	0.0**	0.2***	0.2
US Durable Goods New Orders In	5%	−1080	0.0	0.1***	25%	−1087	0.0	0.2***	0.8***
US Auto Sales Domestic Vehicle	9%	6253	114.6***	0.3***	23%	6203	−156	0***	0***
Adjusted Retail & Food Service	11%	−1426	0.0	0.2***	36%	−1475	0.0	0.2***	1.1***
Adjusted Retail Sales Less Aut	11%	−1533	0.0	0.2***	18%	−1535	0.0	0.3***	0.5
US Durable Goods New Orders To	0%	−1055	0.0**	0.0	15%	−1072	0.0	0.1*	−0.1
GDP US Personal Consumption Ch	4%	−1404	0.0	0.1**	10%	−1400	0.0**	0.1*	−0.2
ISM Non-Manufacturing NMI	4%	457	0.1	0.4**	25%	444	−5.5**	0.4**	1.9*
US Manufacturing & Trade Inven	2%	−2247	0.0	0.1**	14%	−2152	0.0**	0.2***	−0.6**
Philadelphia Fed Business Outl	2%	1738	−0.7	0.3**	12%	1624	7.3***	0.2	−1.1*
MNI Chicago Business Barometer	3%	1368	0.7**	0.3**	6%	1293	2.6	0.4**	0.7
Conference Board US Leading In	6%	−2358	0.0*	0.1***	35%	−2311	0.0**	0.2***	0.2
Conference Board Consumer Conf	14%	1414	0.3	0.7***	20%	1325	0.3	0.6***	0.1
US Empire State Manufacturing	1%	1272	−0.5	0.2	11%	1268	3.0	0.1	1.0
Richmond Federal Reserve Manuf	1%	970	−1.0	0.2	18%	959	−3.2	0.2	1.5**

TABLE 3 (Continued)

Model Statistics/regressors	Anchor-only model R^2	AIC	Intercept	Anchor	Unrestricted model R^2	AIC	Intercept	Anchor	Std
ISM Milwaukee Purchasers Manuf	0%	451	0.0	0.0	12%	456	10.1**	0.1	−0.4
University of Michigan Consume	3%	2109	−0.2*	0.5***	9%	2095	1.1	0.5***	−0.3
Dallas Fed Manufacturing Outlo	8%	652	−3.4***	0.6**	14%	659	0.8	0.6**	−0.7
US PPI Finished Goods Less Foo	9%	−1885	0.0	0.3***	18%	−1734	0.0**	0.3***	0.5
US CPI Urban Consumers MoM SA	1%	−2559	0.0	0.0	21%	−2408	0.0**	0.2***	0.2
US CPI Urban Consumers Less Fo	2%	−2714	0.0	−0.1**	9%	−2523	0.0**	−0.1	−0.3
Bureau of Labor Statistics Emp	1%	−778	0.0	0.0	10%	−718	0.0	0.0	−0.1
US Output Per Hour Nonfarm Bus	3%	−1110	0.0***	0.1**	13%	−1090	0.0	0.1***	0.1
US PPI Finished Goods SA MoM%	13%	−1554	0.0	0.2***	29%	−1533	0.0***	0.4***	1.0**
US Import Price Index by End U	1%	−1661	0.0	0.0	23%	−1677	0.0	0.1***	0.0
US GDP Price Index QoQ SAAR	9%	−1189	0.0	0.2***	41%	−1238	0.0	0.3***	−0.4**
US Personal Consumption Expend	1%	−1660	0.0**	−0.1	27%	−1689	0.0	0.1	−0.6**
US Personal Consumption Expend	1%	−1524	0.0	0.0	12%	−1528	0.0	0.1*	0.1
Average	3.7%	923	-	-	17%	896	-	-	-
Popularity-weighted average	4.0%	110	-	-	17%	102	-	-	-
Average of most popular indicators	4.9%	−312	-	-	18%	−313	-	-	-
% of positive & significant coefficients (P&SC)	-	-	7%	58%	-	-	12%	67%	19%
Popularity-weighted % of P&SC	-	-	6%	64%	-	-	8%	70%	17%
P&SC of most popular indicators	-	-	0%	67%	-	-	11%	67%	22%

Note: The table reports results of anchor-only (restricted) and unrestricted regression models for economic surprises. Regression results are reported per economic indicator. We use Newey–West adjustments to compute coefficient standard errors. The popularity weight provided in the last column of the table uses the sum of our popularity measure across all indicators as base. We measure popularity by averaging the number of analysts that provide forecasts for a given indicator in our sample.

***Significance at the 1% level.

**Significance at the 5% level.

*Significance at the 10% level.

TABLE 3 (Continued)

Model Statistics/regressors	Unrestricted model SurvLag	Skewness	Inflation	Popularity Growth	Stocks	VIX	$\times R^2$ gain	weight
US Initial Jobless Claims SA	0.0	2.1***	582.8	358	−24,468	253***	∞	2.0%
US Employees on Nonfarm Payrol	−0.0001	0.003***	−5	1	−172	−1*	∞	4.4%
U-3 US Unemployment Rate Total	0.0**	1.2***	0.0	0.0	0.0	0.0	∞	4.3%
US Employees on Nonfarm Payrol	0**	2***	−592	1971**	122,767	99	8	0.9%
US Continuing Jobless Claims S	0***	0	0	−2**	−106	1***	4	0.3%
ADP National Employment Report	0	0**	4	−1	506	−1	13	0.9%
US Average Weekly Hours All Em	0.0	2.3***	0.0	0.0	0.6	0.0	∞	0.6%
US Personal Income MoM SA	0.3***	3.3***	0.0	0.0***	0.0	0.0	15	3.4%
ISM Manufacturing PMI SA	0.0	1.1	0.0	0.0	15.4*	0.0	5	3.7%
US Manufacturers New Orders To	0.1	0.1	0.0	0.0	0.0	0.0	7	3.0%
Federal Reserve Consumer Credi	0	0**	−532**	286**	−26,825	7	8	1.9%
Merchant Wholesalers Inventori	0.2	0.9	0.0	0.0	0.0	0.0	9	1.4%
US Industrial Production MOM S	0.4***	3.0***	0.0	0.0	0.0	0.0	3	3.7%
GDP US Chained 2009 Dollars Qo	0.1**	1.7**	0.0	0.0	0.0	0.0	9	5.5%
US Capacity Utilization % of T	0.0	1.9***	0.0	0.0	0.0	0.0*	5	3.2%
US Personal Consumption Expend	0.2***	0.0	0.0	0.0*	0.0	0.0	2	3.5%
US Durable Goods New Orders In	0.4***	2.5***	0.0	0.0***	0.1	0.0	5	3.4%
US Auto Sales Domestic Vehicle	0	0***	−31	8	−5986*	1	3	1.2%
Adjusted Retail & Food Service	0.1	4.2***	0.0	0.0	0.0	0.0***	3	3.1%
Adjusted Retail Sales Less Aut	0.2	2.0**	0.0	0.0	0.0	0.0	2	2.8%
US Durable Goods New Orders To	0.2	1.4*	0.0	0.0***	0.0	0.0	∞	1.4%
GDP US Personal Consumption Ch	0.0	0.2	0.0	0.0	0.0	0.0**	3	0.5%
ISM Non-Manufacturing NMI	0.1***	1.1	−0.1	−0.1**	34.8**	−0.1*	6	1.6%
US Manufacturing & Trade Inven	0.0	1.5***	0.0	0.0	0.0	0.0	7	2.5%
Philadelphia Fed Business Outl	−0.1	2.9***	−0.1	0.0	−18.2	−0.2**	6	2.5%
MNI Chicago Business Barometer	0.0	2.5*	0.1	0.2	−4.2	0.0	2	2.5%
Conference Board US Leading In	0.2***	2.0***	0.0*	0.0**	0.0	0.0	6	2.6%
Conference Board Consumer Conf	0.0	2.6**	0.4*	−0.1	−0.6	−0.1	1	3.3%
US Empire State Manufacturing	−0.1	4.3***	0.2	−0.1	−37.8	−0.3**	11	1.7%
Richmond Federal Reserve Manuf	−0.1	3.3***	0.4	−0.2	−62.7	0.0	18	0.2%

TABLE 3 (Continued)

Model Statistics/regressors	Unrestricted model SurvLag	Skewness	Inflation	Popularity Growth	Stocks	VIX	$\times R^2$ gain	weight
ISM Milwaukee Purchasers Manuf	−0.2**	0.9	0.5	0.5	1.9	0.0	∞	0.1%
University of Michigan Consume	0.0	2.1***	0.2**	−0.1**	9.1	0.0	3	2.6%
Dallas Fed Manufacturing Outlo	0.0	0.3	0.1	−0.3	−52.5	−0.1	2	0.2%
US PPI Finished Goods Less Foo	0.8***	1.1*	0.0	0.0	0.0	0.0	2	3.4%
US CPI Urban Consumers MoM SA	0.2***	1.1***	0.0	0.0	0.0	0.0	21	3.8%
US CPI Urban Consumers Less Fo	−0.4**	0.4	0.0	0.0*	0.0	0.0	5	3.7%
Bureau of Labor Statistics Emp	0.1	0.0	0.0	0.0*	0.0	0.0	10	2.8%
US Output Per Hour Nonfarm Bus	0.1***	0.4	0.0	0.0	0.1	0.0	4	3.0%
US PPI Finished Goods SA MoM%	0.4***	2.0**	0.0	0.0	0.0	0.0	2	3.6%
US Import Price Index by End U	0.3***	2.5***	0.0**	0.0	0.0	0.0*	23	2.0%
US GDP Price Index QoQ SAAR	0.0	3.5***	0.0*	0.0*	0.0	0.0	5	1.1%
US Personal Consumption Expend	−0.1	1.4***	0.0	0.0	0.0	0.0	27	1.3%
US Personal Consumption Expend	0.0	0.6**	0.0	0.0	0.0	0.0	12	0.5%
Average	-	-	-	-	-	-	-	2.3%
Popularity-weighted average	-	-	-	-	-	-	-	-
Average of most popular indicators	-	-	-	-	-	-	-	-
% of positive & significant coefficients (P&SC)	28%	70%	5%	5%	5%	5%	-	-
Popularity-weighted % of P&SC	39%	75%	6%	3%	5%	2%	-	-
P&SC of most popular indicators	33%	78%	11%	0%	11%	0%	-	-

Note: The table reports results of anchor-only (restricted) and unrestricted regression models for economic surprises. Regression results are reported per economic indicator. We use Newey–West adjustments to compute coefficient standard errors. The popularity weight provided in the last column of the table uses the sum of our popularity measure across all indicators as base. We measure popularity by averaging the number of analysts that provide forecasts for a given indicator in our sample.

***Significance at the 1% level.

**Significance at the 5% level.

*Significance at the 10% level.

reach a double-digit level, in contrast with the *anchor-only* model. Most anchor coefficients are statistically significant, like in the *anchor-only* model. In line with earlier results, the *Skewness* coefficients are mostly positive and statistically significant, whereas the coefficient sign is more unstable for the *SurvLag* and *Std* coefficients.

By analyzing individual models' results, we are able to explore an additional aspect of economic indicators, popularity. We measure popularity by averaging the number of analysts that provide forecasts for a given indicator in our sample. In Table 3, popularity is reported in the last column as *Popularity weight*, which uses the sum of our popularity measure across all indicators as the denominator. We also aggregate statistics in Table 3 using the nine most popular US economic indicators employed by Campbell and Sharpe (2009).[11] Overall, we find that model quality is higher for more popular indicators. The R^2 (AIC) weighted using our popularity measure for the *anchor-only* model is 4.0% (110), whereas the (unweighted) average R^2 (AIC) is 3.7% (923). For the *unrestricted* model, the weighted R^2 (weighted AIC) is 17% (102), whereas the average R^2 (average AIC) is 17% (896). Hence, popular indicators seem to be better explained by our explanatory variables. If we compare the percentage of positive and significant coefficients across all models (see bottom of Table 3) to the same measure weighted by popularity and using the most popular indicators, we observe that *ESA* and *Skewness* are more likely to hold with the correct sign among popular indicators. This result applies to the *anchor-only* model and the unrestricted model of *ESA*.

Hence, we conjecture that the strategic and behavioral biases modeled by *Skewness* and *ESA* are more present among popular indicators. This finding makes explicit that the bias in analysis here links to the attention, not to inattention, which is commonly suggested as one of the reasons for behavioral biases in forecasting as argued by Mendenhall (1991), Stickel (1991), Campbell and Sharpe (2009), and Cen et al. (2013).[12] The intuition behind this finding is that as the number of analysts increases, the biases embedded in forecasts are reinforced. In particular, for the case of the *ESA* as the number of forecasters with private (noisy) signals rises, the standard deviation of forecasts decreases. That means that the marginal forecasters (in possession of signals with same quality) must post a forecast in increasingly tighter range to remain close to the consensus, which reinforces the bias. In the case of *Skewness*, as the number of off-consensus forecasts increases with the number of forecasters with superior private information, then the distribution skewness rises, which pushes the marginal forecaster with superior ability to even extremer forecasts. Note that this is a function of the 'winner-take-all contest' where forecasters differentiate from contenders by putting excessive weight on their private signals. Another argument for larger differentiation of forecasts and *Skewness* when the number of forecasters is higher, is that prizes tend to be proportional to the number of forecasters participating in the contest.

3.2 | Popularity effect

We now attempt to more formally capture the effect of popularity on biases represented by the *Skewness* and *ESA* variables. We run panel regression across all US economic indicators using a monthly frequency, with our popularity measure (*Pop*) as an additional predictor. As the economic indicators are expressed in different scales, we normalize them into *Z* scores. In fact, for the sake of comparison, we first run a *Non Pop-effect* model that does not use information on the popularity of indicators. Note that the *Non Pop-effect* model is the panel version of Equation 2. Subsequently, we run Equation 3, a *Pop-effect only* model, that uses our popularity measure (*Pop*) as an additional predictor in the regression. The *ESA interaction* model uses an interaction variable *Pop*ESA* in addition to the *Pop-effect only* model of Equation 4. The *Skewness interaction* model uses an interaction variable *Pop*Skewness* in addition to the *Pop-effect only* model as in Equation 5. Finally, the *Dual interaction* model uses both interactions: *Pop*ESA* and *Pop*Skewness* as in Equation 6.

$$S_t = \alpha + ESA_\varphi + SurvLag_\varphi + Std_\varphi + Skewness_\varphi + Infl_\varphi + Growth_\varphi + Stocks_\varphi + VIX_\varphi + Pop_\varphi + \in_t, \quad (3)$$

$$S_t = \alpha + ESA_\varphi + SurvLag_\varphi + Std_\varphi + Skewness_\varphi + Infl_\varphi + Growth_\varphi + Stocks_\varphi + VIX_\varphi + Pop_\varphi + Pop * ESA + \in_t, \quad (4)$$

$$S_t = \alpha + ESA_\varphi + SurvLag_\varphi + Std_\varphi + Skewness_\varphi + Infl_\varphi + Growth_\varphi + Stocks_\varphi + VIX_\varphi + Pop_\varphi + Pop * Skewness_\varphi + \in_t, \quad (5)$$

$$S_t = \alpha + ESA_\varphi + SurvLag_\varphi + Std_\varphi + Skewness_\varphi + Infl_\varphi + Growth_\varphi + Stocks_\varphi + VIX_\varphi + Pop_\varphi + Pop * ESA_\varphi + Pop * Skewness_\varphi + \in_t, \quad (6)$$

The estimation results of the models above are reported in Table 4. They suggest that the usage of *Pop* as an additional predictor leads to no improvements relative to the original model. However, when *Pop* interacts with either *ESA* and *Skewness*, then it becomes highly significant. The coefficients of the predictors derived from *Pop*

TABLE 4 Popularity effect panel regressions for US economic surprises

	No pop effect	Pop effect only	ESA interaction	Skewness interaction	Dual interaction
Intercept	0.060*	0.056	0.057	0.055	0.055
	(0.034)	(0.042)	(0.042)	(0.042)	(0.042)
ESA	0.159***	0.159***	0.055*	0.159***	0.051
	(0.012)	(0.012)	(0.031)	(0.012)	(0.031)
SurvLag	0.009	0.009	0.012	0.009	0.013
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Std	−0.005	−0.005	−0.005	−0.006	−0.007
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Skewness	0.203***	0.203***	0.203***	0.148***	0.141***
	(0.011)	(0.011)	(0.011)	(0.030)	(0.030)
Infl	0.055***	0.055***	0.057***	0.055***	0.057***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Growth	−0.025*	−0.025*	−0.027**	−0.025*	−0.028**
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Stocks	−0.914*	−0.914*	−0.903**	−0.906*	−0.894**
	(0.850)	(0.851)	(0.850)	(0.850)	(0.850)
VIX	−0.003*	−0.003*	−0.003**	−0.003*	−0.003**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Pop		0.007	0.004	0.007	0.005
		(0.055)	(0.055)	(0.055)	(0.055)
Pop*ESA			0.227***		0.235***
			(0.062)		(0.062)
Pop*Skewness				0.118*	0.132**
				(0.061)	(0.061)
R^2	5.8%	5.8%	5.9%	5.8%	6.0%
Adjusted R^2	5.8%	5.8%	5.9%	5.8%	6.0%
F-stats	65.1	57.9	53.5	52.5	49.1

Note: The table reports results of panel regression models applied for the 43 US economic surprises used in our study. The *Non Pop-effect* model does not use information on the popularity of indicators. The *Pop-effect only* model uses our popularity measure (*Pop*) as an additional predictor in the regression. The *ESA interaction* model uses an interaction variable *Pop*ESA* in addition to the *Pop-effect only* model. The *Skewness interaction* model uses an interaction variable *Pop*Skewness* in addition to the *Pop-effect only* model. The *Dual interaction* model uses both interactions earlier mentioned, *Pop*ESA* and *Pop*Skewness*. We measure popularity by averaging the number of analysts that provide forecasts for a given indicator in our sample. Regression results are reported per economic indicator. We use Newey–West adjustments to compute coefficient standard errors.
***Significance at the 1% level.
**Significance at the 5% level.
*Significance at the 10% level.

are always positive, suggesting that it adds to the effects of the *ESA* and *Skewness* variables when used in isolation. Note that in the *Dual interaction* model, the *ESA* predictor becomes statistically non-significant and drops from 0.159 to 0.051, as the *Pop*ESA* captures the full relation between *ESA* and the target variable. The *Skewness* predictor remains positive and statistically significant despite the introduction of the *Pop*-related variables, indicating its strong relation with economic surprises and an independence from popularity.

The findings above constitute strong evidence that the popularity of economic indicators, measured by the number of analysts posting forecasts, seems to reinforce the presence of forecasting bias. In other words, strategic bias in economic forecasting seems to be more pervasive around popular indicators.

3.3 | Out-of-sample testing

Now that we have identified evidence of a strategic bias and a popularity effect in expected economic surprises, we next attempt to test these findings out-of-sample. The natural benchmark for our out-of-sample analysis is the previously identified anchor-based bias introduced by Campbell and Sharpe (2009), represented by the *ESA* variable. We ultimately want to compare the models specified by Equations 1 and 6. As it is interesting to evaluate how the different components of Equation 6 perform versus the benchmark (Equation 1), we also estimate seven intermediary models in Equations 7a–7g. The anchor-based model with control variables and the *Skewness* variable in Equation 7b equals Equation (2).

$$S_t = \alpha + ESA_\varphi + SurvLag_\varphi + Std_\varphi + Infl_\varphi + Growth_\varphi + Stocks_\varphi + VIX_\varphi + \in_t, \quad (7a)$$

$$S_t = \alpha + ESA_\varphi + SurvLag_\varphi + Std_\varphi + Skewness_\varphi + Infl_\varphi + Growth_\varphi + Stocks_\varphi + VIX_\varphi + \in_t, \quad (7b)$$

$$S_t = \alpha + Skewness_\varphi + \in_t, \quad (7c)$$

$$S_t = \alpha + ESA_\varphi + Skewness_\varphi + \in_t, \quad (7d)$$

$$S_t = \alpha + ESA_\varphi + Pop_\varphi + Pop * ESA_\varphi + Pop * Skewness_\varphi + \in_t, \quad (7e)$$

$$S_t = \alpha + Skewness_\varphi + Pop_\varphi + Pop * ESA_\varphi + Pop * Skewness_\varphi + \in_t, \quad (7f)$$

$$S_t = \alpha + ESA_\varphi + Skewness_\varphi + Pop_\varphi + Pop * ESA_\varphi + Pop * Skewness_\varphi + \in_t, \quad (7g)$$

Note that we have no option but to estimate these models using our panel approach specified when we introduced Equation 6 due to the usage of our popularity measure. We focus on US data only and we evaluate model performance in the last 25% portion of the available history. We estimate models using either the entire data history available prior to the out-of-sample period (i.e., 75% of history, the so-called 75% IS + 25% OOS) or using only the 50% of history prior to the out-of-sample to capture the most recent behavior of the data (i.e., the so-called 50% IS + 25%).[13] We report results in Table 5.

Our first takeaway from Table 5 is that the control variables employed cause a drag to the out-of-sample performance of models when added to more parsimonious models, especially in root mean-square errors (RMSE) terms. In contrast, the popularity measure most of times has a positive or neutral impact to performance. The hit ratio is mostly improved when popularity is added to the *Skewness* model in the 75% IS + 25% OOS run, an impressive improvement of 4.9% in the hit ratio from a low base (49%), as the pure *Skewness* model estimated using the longer sample is the weakest one in the hit-ratio basis. One could think of the popularity measure

TABLE 5 Out-of-sample performance of panel regressions for US economic surprises

		Root mean-square errors (*100)		Hit ratio	
Equation #	Models	75% IS + 25% OOS	50% IS + 25% OOS	75% IS + 25% OOS	50% IS + 25% OOS
1	ESA	1.952	1.953	50.7%	54.6%
7a	ESA + Controls	2.157	2.160	53.1%	51.6%
7b/2	ESA + Controls + Skewness	2.151	2.154	54.1%	51.8%
7c	Skewness	1.962	1.963	49.2%	53.2%
7d	ESA + Skewness	1.949	1.933	51.6%	51.5%
7e	ESA + Pop	1.947	1.949	51.3%	54.1%
7f	Skewness + Pop	1.949	1.948	54.1%	54.8%
7g	ESA + Skewness + Pop	1.946	1.949	52.2%	54.4%
6	ESA + Controls + Skewness + Pop	2.129	2.156	53.4%	51.9%

Note: The table reports out-of-sample performance statistics of panel regressions for economic surprises using US data. The performance statistics are the root mean-square errors (RMSE) and the hit ratio, which evaluates that directional accuracy of the forecast in percentage terms. We ultimately want to compare the models specified by Equations 1 and 6. However, as it is interesting to evaluate how the different components of Equation 6 perform versus the benchmark (Equation 1), we also estimate seven intermediary models as specified by Equations 7a–7g. Note that the anchor-based model with control variables and the *Skewness* variable (Equation 7b) is also specified by Equation (2. Our out-of-sample period corresponds to the last 25% portion of the available history. We estimate models either using the entire data history available prior to the out-of-sample period (i.e., 75% of history), thus named 75% IS + 25% OOS in the table, or using only the 50% of history prior to the out-of-sample, thus named 50% IS + 25% OOS.

adding more value when the earlier part of the sample is used for estimation, as popular indicators may correlate with indicators with higher longevity. This would happen because popularity would put more weight into better performing forecasts of long-standing indicators (where biases are likely to be higher) versus weaker performing forecasts of new indicators in the earlier part of the sample relative to its later part. Popularity adds substantial value to the *ESA+Skewness* as well, nearly 3% in the hit ratio in the 50% IS + 25% OOS run. This model is most powerful in a statistical sense; that is, it has the lowest RMSE, but it ranks among the worst ones in an economic sense, as hit ratios are mostly below 52%. Once popularity is added, this model becomes a top performer also on a hit-ratio basis. For the *ESA* model, the benchmark model, the usage of the popularity measure seems to positively impact performance on a RMSE basis, though, the result is mixed for the hit ratio. The benchmark model is mediocre in RMSE basis and the hit ratio using the earlier part of the sample but strong when only the last part of the sample is used for estimation. When we directly compare the benchmark model with our suggested *Skewness* model, we observe that, in isolation, the *ESA* is a stronger model. However, when popularity is added then the *Skewness* model becomes superior and the best model on an average hit-ratio basis across the two samples used. Adding *Skewness* to *ESA* has an average positive impact of 0.6% in the hit ratio provided that popularity is present, suggesting that our two innovations add value to the current benchmark model.

Our out-of-sample results confirm our findings observed in-sample as they reiterate the presence of the strategic bias reflected by the *Skewness* measure and popularity effect in the forecasting of economic surprises.

4 | ROBUSTNESS TESTS

4.1 | Economic surprise models across regions

As a robustness test, we apply Equations 1 and 2 across continental Europe, the United Kingdom, and Japan using, respectively, 147, 37, and 35 indicators.[14] Table 6 indicates that our results for these three regions are qualitatively the same as the ones reported for the United States. The *unrestricted* models tend to improve the R^2 of *anchor-only* ones. The coefficients for the *ESA* and *Skewness* factors are as expected mostly positive. These two coefficients are positive between 56% and 75% of all times, which is, however, lower than the percentage of correct signs found for the United States. Yet, among coefficients for all factors (including control variables), *ESA* and *Skewness* remain the ones that are mostly positive. Moreover, in terms of statistical significance, the specified models (*anchor-only* and *unrestricted*) for Europe, Japan, and the United Kingdom perform worse than the US model, as the percentage of coefficients that are significant is, in general, lower than for the United States.

We conjecture that the difference in presence of biases in economic forecasting across the different regions may be explained by the number of experts dedicated to economic forecasts across these countries. The average number of analysts providing forecasts across all indicators of our sample is 44 for the United States. For Europe, Japan, and the United Kingdom this number is 9, 13, and 15, respectively. We argue that, as the number of forecasters increases for a specific indicator or within a country, it becomes more likely that (1) convergence towards the previous release happens simply by the law of large numbers; (2) some forecasters possess superior private information; and (3) such private information is revealed by the skewness in forecasts, given strategic behavior by experts.

4.2 | Economic surprise models through time

We investigate in this section how the strong relations found between economic surprises and the *ESA* and *Skewness* factors in our main analysis behave over time. To perform a stability test we employ a panel regression version of Equation 2, our main predictive model for economic surprises. We use a panel regression[15] because, as we mostly use monthly data, rolling regressions using a single indicator would hardly contain enough observations to capture statistically significant links between surprises and explanatory variables. Given that economic surprises and some explanatory variables, such as *ESA* are expressed in different scales, we normalize them into Z scores.

Our results are reported in Figure 1, which depicts the coefficient values and p values of the predictors *ESA*, *Skewness*, *Std*, and *SurvLag* over time. We observe that the coefficients for the variables *ESA* and *Skewness* are positive and statistically significant, with a few exceptions. At the same time, the signs of the coefficients for *Std* and *SurvLag* are unstable, fluctuating between positive and negative, beyond being mostly statistically insignificant. Additionally, we observe that at the start of our sample the magnitude of the *ESA* coefficient is more than twice that of the coefficient of *Skewness*, indicating a larger economic significance of the normalized *ESA* in estimating economic surprises. Nevertheless, the magnitude of the coefficient for *Skewness* steadily increases through our sample and after 2009 it becomes structurally larger than the one for *ESA*. This behavior reiterates our main findings, which

TABLE 6 Aggregated results of anchor-only and unrestricted economic surprise models for Europe, the United Kingdom, and Japan

Region	Continental Europe		United Kingdom		Japan	
Model	Anchor-only	Unrestricted	Anchor-only	Unrestricted	Anchor-only	Unrestricted
Panel A—Percentage of statistical significance per factor						
Intercept	0.22	0.29	0.33	0.19	0.24	0.16
Bias	0.44	0.37	0.33	0.31	0.42	0.28
Std		0.18		0.42		0.16
SurvLag		0.20		0.33		0.31
Skewness		0.09		0.44		0.13
Infl		0.11		0.11		0.09
Growth		0.14		0.19		0.13
Stocks		0.15		0.22		0.13
VIX		0.29		0.25		0.06
Panel B—Explanatory power (R^2)						
Mean R^2	8%	25%	3%	18%	3%	16%
Median R^2	3%	16%	1%	13%	2%	10%
Stdev R^2	17%	24%	5%	13%	3%	20%
Panel C—Percentage of positive coefficients						
Intercept	0.37	0.53	0.56	0.50	0.52	0.78
Bias	0.66	0.62	0.58	0.72	0.70	0.69
Std		0.42		0.56		0.44
SurvLag		0.47		0.36		0.31
Skewness		0.56		0.75		0.59
Infl		0.40		0.25		0.44
Growth		0.44		0.44		0.34
Stocks		0.48		0.44		0.47
VIX		0.26		0.42		0.19

Note: Panel A reports the percentage of statistical significant coefficients (factors) across anchor-only and unrestricted regression models for economic surprises of economic indicators for Europe, the United Kingdom, and Japan. For example, 0.44 found for the *ESA* factor within the anchor-only model for Europe means that 44% of such *ESA* factor across the individual regressions run for the European economic indicators are statistically significant at the 10% level. Panel B reports the mean, median, and standard deviation of the explanatory power (R^2) achieved across all indicator-specific regressions. Panel C reports the percentage of positive coefficients across anchor-only and unrestricted regression models for economic surprises of the same economic indicators.

suggest that both *ESA* and *Skewness* are strongly connected to economic surprises. It also indicates that the *Skewness* factor has gained relevance lately, whereas the impact of our anchor-based factor (*ESA*) has diminished over the sample period, which may be linked to the seminal publication of Campbell and Sharpe (2009).

4.3 | Biases by individual forecasters

In this section we evaluate if the anchor and the strategic bias found in our main experiments are also observed in data from individual forecasters. The availability of data per forecaster is much less than aggregated forecasters' data, therefore, our analysis uses only three economic indicators and ranges from 2015 to 2018 only. The US economic indicators investigated are the Consumer Price Index (CPI) headline inflation, the unemployment rate, and the quarterly annualized GDP. The forecasters' data are downloaded from Bloomberg's FRCS function, which compiles forecasts for multiple end-of-quarter numbers of some economic indicators. Therefore, this analysis differs from the main empirical analysis in this paper, which evaluates forecasts and economic releases at the frequency of releases' publication, where the coming release number is always the one in focus. Differently, here we evaluate forecasts on a daily or monthly basis but the economic release in focus is always the end-of-quarter number. Thus, despite the fact that the CPI inflation and the unemployment rate are published on a monthly basis, the

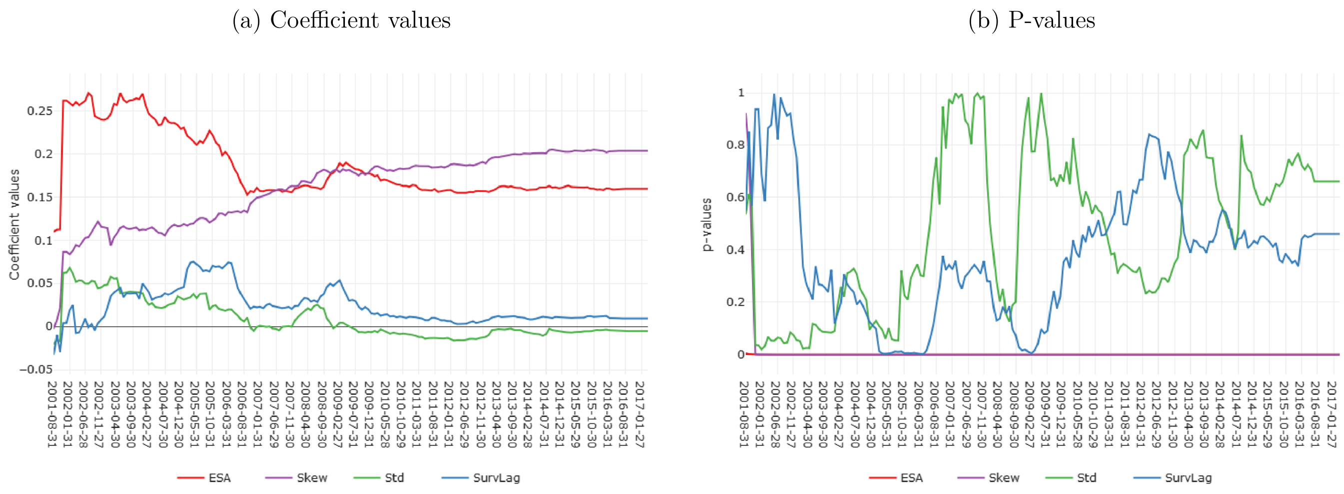

FIGURE 1 Economic surprise models through time. The line plots above depict the coefficient values and p value of predictors *ESA*, *Skewness*, *Std*, and *SurvLag* through time, respectively in panels (a) and (b). The coefficient of *ESA* and *Skewness* are positive and statistically significant, with few exceptions, whereas the coefficients for *Std* and *SurvLag* fluctuate between a positive and a negative sign, beyond being mostly not statistically significant [Colour figure can be viewed at wileyonlinelibrary.com]

forecasts used may target 1- to 3-month ahead releases. As forecasts can be adjusted on a continuous basis, our data set comprises a set of daily forecasts per forecasting professional, whereas realized numbers are constant in the interval between releases. This way, the median forecast changes between releases, whereas the anchor (the previous release) is kept constant.

The panel consists of 98 forecasters within the period specified. Nevertheless, to avoid an unbalanced panel, we exclude forecasters who failed to submit a forecast to any of the targeted quarter numbers. This adjustment leaves us with 66 forecasts for the headline CPI, 62 for the unemployment rate, and 77 forecasts for the GDP for the 14 quarters evaluated.

First, to assess whether forecasts use the previous release as an anchor, we estimate the following model, which resembles Equation A3:

$$F_t = \alpha_f + A_{t,f} + \in_t, \tag{8}$$

where $A_{t,f}$ is the last available economic release and F_t is the most recent forecast.

Second, we check if the *ESA* is capable of forecasting economic surprises. Thus, we estimate model Equation 9, which is the panel version of model Equation 1, as follows:

$$S_t = \alpha_f + ESA_{t,f} + \in_{t,f}, \tag{9}$$

where $ESA_{t,f}$ is computed per individual forecaster, $f = 1 \ldots 66$ for the US CPI inflation case, $f = 1 \ldots 62$ for the employment rate, and $f = 1 \ldots 77$ for the GDP.

Third, we check if the individual forecaster *Skewness*$_{t,f}$, namely, the difference between the individual forecaster estimate and the median estimate, $F_{t,f} - F_t$, can also help to forecast economic surprises, using Equation 10:

$$S_t = \alpha_f + ESA_{t,f} + Skewness_{t,f} + \in_{t,f}. \tag{10}$$

The estimation results of Equations 8–10 are reported in Table 7. As we run regressions using forecasts at the daily and monthly frequency, we report results separately in Panels A and B.

A first observation from Table 7 is that the previous release anchor remains strongly connected to individual forecasters' expectations, especially for inflation and for unemployment. The previous release explains up to 44% and 82% of the variation of inflation and unemployment rate forecasts across the panel. For GDP, this explanation is much smaller, at around 4%. These findings are homogeneous across the two data frequencies used.

Further, *ESA* is also found to be statistically significantly linked to economic surprises, explaining from 7% (for unemployment) to 51% (for inflation). In contrast to our main results, *ESA* is negatively linked to surprises across the three indicators and the two frequencies evaluated. More importantly, *Skewness* again boosts the explanatory power of regressions when added to *ESA* to explain economic surprises. The explanatory power of inflation surprises rises approximately 25 percentage points when *Skewness* is used. In the case of unemployment, the explanatory power of regressions increases from 7% to 21%. For GDP, the explanatory power also rises but to a lesser extent, roughly 5% across the two frequencies used. The AIC also indicates that the quality of fit improves across regressions. In line with the 'signaling hypothesis', the sign of *Skewness* is positive across all

TABLE 7 Anchor and rational biases by individual forecasters

Panel A—Daily data							Panel B—Monthly data					
US CPI inflation												
Explained variable	**R^2**	**AIC**	**Anchor**	**ESA**	**Skewness**	**Std**	**R^2**	**AIC**	**Anchor**	**ESA**	**Skew**	**Std**
Individual forecasts (Fi,t)	38%	3589	0.477***				44%	3265	0.5388***			
Economic surprise (Si,t)	51%	1480		−0.539***			41%	763		−0.430***		
Economic surprise (Si,t)	75%	−558		−0.798***	0.786***		66%	−839		−0.701***	0.687***	
Economic surprise (Si,t)	75%	−575		−0.781***	0.769***	−0.367***	66%	−841		−0.699***	0.685***	−0.052
US unemployment rate												
Explained variable	**R^2**	**AIC**	**Anchor**	**ESA**	**Skewness**	**Std**	**R^2**	**AIC**	**Anchor**	**ESA**	**Skew**	**Std**
Individual forecasts (Fi,t)	82%	−3216	0.774***				82%	−3169	0.768***			
Economic surprise (Si,t)	7%	−3863		−0.177***			7%	−3716		−0.182***		
Economic surprise (Si,t)	21%	−4343		−0.598***	0.598***		21%	−4182		−0.600***	0.601***	
Economic surprise (Si,t)	21%	−4348		−0.619***	0.613***	0.063	21%	−4186		−0.618***	0.613***	0.089
US annualized GDP												
Explained variable	**R^2**	**AIC**	**Anchor**	**ESA**	**Skewness**	**Std**	**R^2**	**AIC**	**Anchor**	**ESA**	**Skew**	**Std**
Individual forecasts (Fi,t)	4%	2289	0.208***				4%	2170	0.213***			
Economic surprise (Si,t)	17%	1996		−0.306***			13%	1776		−0.254***		
Economic surprise (Si,t)	23%	1917		−0.398***	0.384***		18%	1723		−0.331***	0.316***	
Economic surprise (Si,t)	52%	1410		−0.764***	0.747***	5.282***	43%	1375		−0.704***	0.687***	4.828***

Note: This table reports results of panel regressions given by Equations 8–10 for the US CPI headline inflation, the US unemployment rate, and the quarterly US annualized GDP. Equation 8 tests for the presence of anchor in individual economic forecasts, whereas Equations 9 and 10 attempt to predict economic surprises using *ESA* and *Skewness* from individual forecasters' data. Panel A reports regression results when forecasts are allowed to change on a daily changes, whereas Panel B reports regressions when forecasts are monthly.

***Significance at the 1% level.

**Significance at the 5% level.

*Significance at the 10% level.

regressions and all coefficients are statistically significant. These results are consistent across the two frequencies evaluated and strengthen our main findings.

4.4 | Popularity effect for Europe, the United Kingdom, and Japan

In the following we check for the presence of the popularity effect found for the United States (see Section 3.3) in Europe, the United Kingdom, and Japan. We apply the panel regression approach of Equation 5 across these different regions. To perform an out-of-sample test of the popularity effect found for the United States, we also use the model parameterized with the US data to forecast economic surprises in the other geographies. The estimates for Europe, the United Kingdom, and Japan are reported in Table 8 below. For comparison purposes, we also report the model previously estimated for the United States in Table 8.

TABLE 8 Popularity effect pooled regressions for the United States, Europe, the United Kingdom, and Japan

	United States	Europe	United Kingdom	Japan
Intercept	0.055	0.036**	−0.033	0.039
	(0.042)	(0.018)	(0.041)	(0.038)
ESA	0.051	−0.164***	−0.053	−0.228***
	(0.031)	(0.010)	(0.035)	(0.022)
SurvLag	0.013	−0.059***	−0.084***	−0.082***
	(0.013)	(0.007)	(0.015)	(0.015)
Std	−0.007	−0.051***	−0.041***	0.000
	(0.012)	(0.007)	(0.014)	(0.014)
Skewness	0.141***	−0.100***	0.116**	0.034
	(0.030)	(0.022)	(0.046)	(0.024)
Infl	0.057***	−0.013**	−0.027**	0.023*
	(0.011)	(0.006)	(0.012)	(0.012)
Growth	−0.028**	0.041***	−0.008	0.008
	(0.014)	(0.007)	(0.013)	(0.014)
Stocks	−0.894**	−0.222***	1.055	0.576
	(0.850)	(0.480)	(0.862)	(0.954)
VIX	−0.003**	−0.002***	0.003	−0.002
	(0.002)	(0.001)	(0.002)	(0.002)
Pop	0.005	0.000	−0.144	0.010
	(0.055)	(0.060)	(0.138)	(0.109)
Pop*ESA	0.235***	1.265***	0.429**	1.138***
	(0.062)	(0.072)	(0.191)	(0.123)
Pop*Skewness	0.132**	0.838***	−0.367	0.577***
	(0.061)	(0.136)	(0.298)	(0.131)
R^2	6.0%	2.0%	1.1%	3.5%
Adjusted R^2	6.0%	2.0%	1.1%	3.5%
F-stats	49.1	55.0	7.1	23.7

Note: The table reports results of pooled regression models applied for the United States, Europe, the United Kingdom, and Japan. The *Non Pop-effect* model does not use information on the popularity of indicators. The *Pop-effect only* model uses our popularity measure (*Pop*) as an additional predictor in the regression. The *ESA interaction* model uses an interaction variable *Pop*ESA* in addition to the *Pop-effect only* model. The *Skewness interaction model uses an interaction variable Pop*Skewness* in addition to the *Pop-effect only* model. The *Dual interaction* model uses both interactions earlier mentioned, *Pop*ESA* and *Pop*Skewness*. We measure popularity by averaging the number of analysts that provide forecasts for a given indicator in our sample. Regression results are reported per economic indicator. We use Newey–West adjustments to compute coefficient standard errors.

***Significance at the 1% level.

**Significance at the 5% level.

*Significance at the 10% level.

In line with the results found for the United States, the interaction variables *Pop*ESA* and *Pop*Skewness* are mostly statistically significant and positive. Whenever these variables have a negative sign, they are not statistically significant. We find that *Pop*ESA* is always statistically significant and positive, whereas *ESA* has a negative coefficient when statistically significant. This finding suggests that the original positive impact of *ESA* in the economic surprises, captured by Equation 2, is now fully captured by the interaction variable, a shift that was also observed for the United States when different panel regressions were compared in Section (3.3). In contrast to the United States, when *Pop*Skewness* is positive and statistically significant, *Skewness* is either negative or not statistically significant, suggesting that outside the United States the influence of *Skewness* is either channeled via the interaction term or *Skewness* in isolation, not via the two channels. For instance, for the United Kingdom the link between *Skewness* and economic surprises is independent of popularity, whereas for Europe and Japan the interaction term *Pop*Skewness* matters more than *Skewness* in isolation. Similar to the United States, *Pop* is never statistically significant across the other geographies. In summary, the results above suggest that popularity effect reported for the United States is to a great extent also found in Europe, the United Kingdom, and Japan as well, more strongly so via *ESA* but are also via *Skewness* in Europe and Japan. This result is in line with expectations as economic data in the United States is likely more closely followed than similar data in any other country or region. Put differently, if 'relative popularity' across the different economic indicators is a catalyst for bias in economic forecasts, it is expected that such biases are more prevalent where 'absolute popularity' is higher, that is, in the geography with a higher number of forecasters. This expectation is based on the notion that, as the number of biased forecasters increases, the biases tend to become stronger. Evidence for these biases being stronger in the United States is given by the average number of forecasters per indicator per country, which is 44 for the United States, compared to only 9 for Europe, 15 for the United Kingdom, and 14 for Japan.

To test the popularity effect model out-of-sample, we calculate the RMSE produced by the model specified in Equation 5 parameterized with US data, to forecast economic surprises in Europe, United Kingdom, and Japan. The benchmark for the RMSE calculated for this out-of-sample test is the in-sample RMSE produced by the same model when estimated with data from the respective countries, that is, the models presented in Table 8 for Europe, United Kingdom, and Japan. Out-of-sample RMSEs (×100) are 0.586, 1.144 and 1.203 for Europe, United Kingdom, and Japan, whereas in-sample RMSEs are, respectively, 0.578, 1.113 and 1.184.[16] Hence, the increase in RMSE of the out-of-sample model versus the in-sample models is only 2.8% at the highest, which we consider low when comparing an out-of-sample statistic to an in-sample one.

A final out-of-sample test is performed by comparing out-of-sample RMSEs of the model parameterized using 75% of US data and evaluated using the remaining 25% of the data in the other geographies. The benchmarks in this case are the models parameterized using 75% of the data from Europe, United Kingdom, and Japan, respectively. In this case, as all RMSEs are out-of-sample, the gap in RMSEs is much lower. RMSEs for Europe, United Kingdom, and Japan, when the model is estimated with US data are 1.375, 2.842, and 2.752, whereas RMSEs when the model is estimated with the country-specific data are 1.359, 2.826, and 2.712. The RMSE increase delivered by the United States-based model equals 1.5% at the highest, which we again consider a low number. We thus note that the popularity effect observed in the United States also holds in Europe, United Kingdom, and Japan.

5 | CONCLUSION

This paper provides evidence that the information contained in the distribution of economic forecasts from surveys is an additional source of information in predicting economic surprises across a wide range of indicators. In particular, we argue that the skewness in the distribution of economic forecasts reflects a strategic bias and contains information, as proposed by the literature on strategic behavior of forecasters. According to this stream of research, forecasters have dual and contradicting objectives, that is, forecast accuracy and publicity. Forecasters often stay close to the 'pack' (and exhibit herding behavior) to avoid being wrong. Eventually, though, when in the possession of what they perceive to be superior private information, they signal confidence in their own forecasts by issuing off-consensus forecasts, giving rise to a skewed distribution of forecasts. Thus, by using information from these controversial forecasts, predictability of economic surprises is improved versus the usage of the anchor bias, a well-documented behavioral bias in economic forecasting. This strong finding is confirmed by us empirically, through both in- and out-of-sample analysis.

The strong link between economic surprises and skewness in the distribution of forecasts also holds in the data from individual forecasters and for continental Europe, the United Kingdom, and Japan, however, to a lesser extent than in the United States. We find that the importance of the skewness in the distribution of economic forecasts in predicting economic surprises has been steadily increasing through time and versus the anchor bias.

Predictability of economic surprises is found to be stronger for popular indicators in the United States. When we move from widely followed US indicators towards less watched ones, the strategic bias documented becomes less pervasive: a popularity effect in the prediction of economic surprises. This popularity effect also holds in continental Europe, the United Kingdom, and Japan but, again, in a weaker form than in the United States as the number of forecasters per indicator in these regions is much smaller than in the United States. The popularity effect is particularly strong in boosting performance of models based in the skewness in the distribution of forecasts but also present in anchor-based models.

There are three key implications of our findings. First, they enable a better understanding of the informational content of the skewness in the distribution of economic forecasts by regulators, policy makers, and market participants. Second, our strong results provide evidence that the anchor bias is not the single bias widely found in economic forecasting. As a result, benchmarks for assessment of economic surprise models should be complemented with our suggested skewness measure. Third, the popularity effect identified supports the usage of a weighted scheme versus an unweighted one in the construction of economic surprise indexes. The popularity effect also reinforces the attention of economist and market participants to popular economic indicators.

ACKNOWLEDGMENTS

We thank the two anonymous referees, Prof. Massimiliano Marcellino (the Editor), and seminar participants at the Econometrics and Financial Data Science workshop of the Henley Business School in Reading, the APG Asset Management Quant Roundtable, the MAN-AHL Research Seminar in 2017, the 2018 Annual Conference of the Swiss Society for Financial Market Research (SGF) Conference in Zurich, the 2018 EEA-ESEM Conference in Cologne, the 2018 European Finance Association (EFA) in Warsaw, and the 2018 Research in Behavioral Finance Conference (RBFC) in Amsterdam for their helpful comments. We thank APG Asset Management for making available part of the data.

DATA AVAILABILITY STATEMENT

The data on economic releases, which support the findings of this study is available on request.

ORCID

Luiz Félix https://orcid.org/0000-0003-0464-0585

ENDNOTES

[1] It attempts to explain why earnings estimates by equity analysts are systematically overoptimistic. De Bondt and Thaler (1990) suggest that equity analysts suffer from a cognitive failure which leads them to overreact and have too extreme expectations. Mendenhall (1991) argues that underreaction to past quarterly earnings and stock returns contributes to an overoptimistic bias in earnings.

[2] A later branch of the literature proposes that biases are caused by a strategic bias. Michaely and Womack (1999) advocate that equity analysts often recommend companies that their employer has recently taken public. In the same vein, Tim (2001) suggests that a strategic bias exists within corporate earnings forecasts because analysts trade off this bias to improve management access and forecast accuracy.

[3] Evidence that economic forecasting is a "winner-take-all contest" is that analyst awards, such as the StarMine Analyst Awards from Refinitiv, a data provider, recognize the world's top individual on the different categories of the award. See https://www.analystawards.com/awards.php?t=2.

[4] Anchoring in forecasting seems not to be, however, restricted to economic data releases. Cen et al. (2013) show that anchoring also plays a role in FEPS by stock analysts, who tend to issue optimistic (pessimistic) forecasts when FEPS are lower (higher) than the median.

[5] The ESA_t variable builds fully on the work of Campbell and Sharpe (2009). The only difference between our approach and theirs is that they model the anchor as the average value of the forecasted series over a number of previous releases, whereas our anchor variable relies only on the previous release. A generic formulation of research applied to forecast bias as well as a derivation of the ESA_t variable is provided in Appendix A.

[6] Note that we use *ESA* variable of Campbell and Sharpe (2009) instead of a variable motivated by a strategic bias with similar hypothetical implication (i.e., concentration in forecasts) because *ESA* is an empirical test, whereas limited empirical evidence has been produced over a connected strategic bias.

[7] PCA is an unsupervised machine learning method that transforms correlated variables into a set of orthogonal variables, so-called principal components.

[8] Beber et al. (2015) split indicators into four categories, that is, output, employment, sentiment, and inflation. We aggregate output, employment, and sentiment indicators into the single category growth. As our set of indicators perfectly matches the one of Beber et al. (2015), this attribution exercise is straightforward. The only different nuance is that Beber et al. (2015) adjust series using 1- and 12-month changes, whereas we use 6-month changes across all non-stationary indicators.

[9] The coefficient γ of Equations 1 and A5 is excluded from this model and subsequent ones for conciseness of presentation. We use subscript ϕ (i.e., $t-1$) to state that the model is predictive. In reality, the subscript t still suggests a prediction as most economic indicator surveys close for forecast submission days before

the release. In Bloomberg, surveys close one business day prior to the announcement.

[10] We also apply Equations 1 and (2 where the predictor *ESA* (i.e., $F_t - A$) is not calculated relative to the median forecast but to the mean forecast. The rationale behind this robustness check is to identify whether or not the median forecast is an inefficient predictor of economic surprises (versus the mean) and to test whether the predictive power of our *Skewness* measure vanishes through the use of the mean forecast within *ESA*. We find that our outcomes change only marginally, thus leaving our main results unaffected.

[11] The indicators used by Campbell and Sharpe (2009) are the NFP Employment Indicator, Michigan Consumer Confidence, Consumer Price Index (CPI) headline and Core, Industrial Production, ISM Manufacturing Index and Retail Sales Headline and ex-Autos. New Homes Sales is also used by these authors but as housing data is out-of-scope of our set of economic indicators, it is not part of our data set.

[12] We see no contradiction between attention seeming to reinforce bias in the current case while dampening biases in other cases. Inattention in the case of behavioral biases typically applies to the processing of information; thus, it is the root cause of the bias. In the case of the strategic bias we focus here where publicity and public endorsement is sought (driven by incentives); attention is ultimately dictated by the pubic or forecasters' clients, who focus on the indicators they value the most. Thus, attention of forecasters is not the root cause of the bias but simply follows from the popularity of one indicator versus others, driven by the forecasters' clientele attention. This differentiation is the main reason why we do not refer to the effect investigated here as an attention effect but, rather, a popularity effect, which we further explore in the coming section.

[13] We estimate models using the most recent data only as we observe a clear shift in coefficient size for *ESA* and *Skewness* as reported in the upcoming Section 4.2 and Figure 1.

[14] The overview of economic news releases for these regions is available upon request.

[15] We use a fixed-effect panel regression model. As an alternative method, we test a pooling regression model. This method delivers results that are qualitatively the same.

[16] Note that RMSE is calculated over the entire history as the test is out-of-sample relative to the country data used for estimation, not relative to a specific part of the history used for estimation.

[17] For simplicity we omit the intercept term of this and the following regression.

REFERENCES

Aggarwal, R., Mohanty, S., & Song, F. (1995). Are survey forecasts of macroeconomic variables rational? *Journal of Business*, *68*, 99–119.

Ashiya, M. (2006). Testing the rationality of forecast revisions made by the IMF and the OECD: *Journal of Forecasting*, *25*, 25–36.

Ashiya, M. (2009). Strategic bias and professional affiliations of macroeconomic forecasters. *Journal of Forecasting*, *28*, 120–130.

Ashiya, M., & Doi, T. (2001). Herd behavior of Japanese economists. *Journal of Economic Behavior & Organization*, *46*, 343–346.

Beber, A., Brandt, M., & Luisi, M. (2015). Distilling the macroeconomic news flow. *Journal of Financial Economics*, *117*, 489–507.

Campbell, S., & Sharpe, S. (2009). Anchoring bias in consensus forecasts and its effect on market prices. *Journal of Financial and Quantitative Analysis*, *44*, 369–390.

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

Cen, L., Hilary, G., & Wei, K. (2013). The role of anchoring bias in equity market: evidence from analysts' earnings forecasts and stock returns. *Journal of Financial and Quantitative Analysis*, *48*, 47–76.

De Bondt, W., & Thaler, R. (1990). Do security analysts overreact? *American Economic Review*, *80*, 52–57.

Ehrbeck, T., & Waldmann, R. (1996). Why are professional forecasters biased? Agency versus behavioral explanations. *Quarterly Journal of Economics*, *111*, 21–40.

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

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

Legerstee, R., & Franses, P. (2015). Does disagreement amongst forecasters have predictive value? *Journal of Forecasting*, *34*, 290–302.

Mankiw, N., Reis, R., & Wolfers, J. (2003). Disagreement about inflation expectations. *NBER Macroeconomics Annual*, *18*, 209–248.

Mendenhall, R. (1991). Evidence of possible underweighting of earnings-related information. *Journal of Accounting Research*, *29*, 140–170.

Michaely, R., & Womack, K. (1999). Conflict of interest and the credibility of underwriter analyst recommendations. *Review of Financial Studies*, *12*, 653–686.

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

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

Schirm, D. (2003). A comparative analysis of the rationality of consensus forecasts of U.S. economic indicators. *Journal of Business*, *76*, 547–561.

Stickel, S. (1991). Common stock returns surrounding earnings forecast revisions: more puzzling evidence. *The Accounting Review*, *66*, 402–416.

Tim, T. (2001). Rationality and analysts' forecast bias. *Journal of Finance*, *61*, 369–385.

Tversky, A., & Kahneman, D. (1974). Judgement under uncertainty: heuristics and biases. *Science*, *185*, 1124–1131.

Zhang, X. (2006). Information uncertainty and analyst forecast behavior. *Contemporary Accounting Research*, *23*, 565–590.

AUTHOR BIOGRAPHIES

Luiz Felix is the expert portfolio manager at APG Asset Management. Luiz holds a PhD in Finance from the VU University Amsterdam (NL) and a MSc in Finance and Investments from Durham University (UK) and has published papers in behavioral and quantitative finance in reputable journals. His latest research applies artificial intelligence techniques to investing, allowing him to develop natural language processing- and machine learning-based investment strategies. Luiz works in financial markets as a quantitative portfolio manager and researcher since 2001. Previously, he worked for ABP Investments and ABN AMRO Asset Management among other institutions.

Roman Kräussl is a Professor of Finance at the University of Luxembourg and Visiting Fellow at the Hoover Institution at Stanford University. He studied economics at the University of Bielefeld and got his PhD in Financial Economics at the Center for Financial Studies (CFS) at Goethe-University, Frankfurt. As Head of Quantitative Research at Cognitrend, a spin-off from Deutsche Bank specialized in behavioral finance and the development of mathematical trading models, he was closely involved with the financial industry. His research focuses on alternative investments, including private equity and infrastructure.

Philip Stork is part-time full professor at Vrije Universiteit Amsterdam. He completed his MSc in Econometrics at Erasmus University Rotterdam, where he also obtained his PhD in 1994 through the Tinbergen Institute. He worked for MeesPierson, Fortis Bank, and Van der Hoop Bankiers as Head of Derivatives, Managing Director Investment Banking, and Executive Director. In the past, he worked for the European Parliament, Banking Code Monitoring Committee, IMC Pacific in Sydney, Massey University in Auckland, and various financial institutions. He currently is a member of the board of Trustees at ABP and at Pension fund Vervoer. He publishes actively on the behavior of financial markets and the development of investment strategies.

How to cite this article: Félix L, Kräussl R, Stork P. Strategic bias and popularity effect in the prediction of economic surprises. *Journal of Forecasting*. 2021;40:1095–1117. https://doi.org/10.1002/for.2764

APPENDIX A: FORECAST BIASES AND ANCHORING AND RATIONALITY TESTS

This appendix introduces the generic formulation of research applied to forecast biases, based on Aggarwal et al. (1995), Schirm (2003), and Campbell and Sharpe (2009). A rationality test of economic forecasts is run by regressing the actual release, A_t, as the explained variable, on the most recent forecast, F_t, as the explanatory variable:

$$A_t = \beta_1 F_t + \in_t. \tag{A1}$$

Rationality holds when β_1 is not significantly different from one, while β_1 significantly higher (lower) than one suggests a structural downward (upward) bias of forecasts.[17] Observing serial correlation in the error term suggests irrationality, as one would be able to forecast A_t using an autoregressive model.

A more intuitive rationality test, suggested by Campbell and Sharpe (2009), is achieved by subtracting the forecast from the left side of Equation A1:

$$S_t \equiv A_f + F_t = \beta_2 F_t + \in_t, \tag{A2}$$

where the new explained variable is the forecast error or ‘economic surprise’, S_t. In Equation A2, rationality holds when β_2 is not significantly different from zero; otherwise, a structural bias is perceived. For the case of anchoring, we dissect the forecast bias using

$$F_t = \lambda E[A_t] + (1-\lambda)A, \tag{A3}$$

where $E[A_t]$ is the forecaster's unbiased prediction, and A is the anchor, which equals the value of the previous release of the indicator. If $\lambda < 1$, the consensus forecast is anchored to the previous release. If $\lambda = 1$, no anchor is observed. By applying expectations to Equation A2, then, substituting $E[A_t] = E[S_t] + F_t$ into Equation A3, we obtain after some manipulations Equation A4c:

$$F_t = \lambda(E[S_t] + F_t) + (1-\lambda)A, \tag{A4a}$$

$$\lambda E[S_t] = F_t - \lambda F_t - A + \lambda A, \tag{A4b}$$

$$E[S_t] = \frac{(1-\lambda)(F_t - A)}{\lambda}, \tag{A4c}$$

introducing $\gamma \frac{(1-\lambda)}{\lambda}$ and unveiling the intercept (α), we find

$$S_t = \alpha + \gamma ESA_t + \in_t, \tag{A5}$$

which enables a direct test of anchoring, if $\gamma > 0$, where ESA_t ($\equiv F_t - A$) is the expected surprise conditional to an anchor.