CENTER FOR FINANCIAL STUDIES

CFS Working Paper Series

No. 596

Luiz Félix, Roman Kräussl, and Philip Stork

Predictable Biases in Macroeconomic Forecasts and Their Impact Across Asset Classes

Center for Financial Studies | Goethe University | Theodor-W.-Adorno-Platz 3 | 60323 Frankfurt am Main | Germany
Phone +49 69 798 30050 | Fax +49 69 798 30077 | info@ifk-cfs.de | www.ifk-cfs.de

The CFS Working Paper Series

Center for Financial Studies | Goethe University | Theodor-W.-Adorno-Platz 3 | 60323 Frankfurt am Main | Germany
Phone +49 69 798 30050 | Fax +49 69 798 30077 | info@ifk-cfs.de | www.ifk-cfs.de

Predictable Biases in Macroeconomic Forecasts and Their Impact Across Asset Classes*

Luiz Félix
VU University Amsterdam
APG Asset Management

Roman Kräussl
Luxembourg School of Finance,
University of Luxembourg

Philip Stork†
VU University Amsterdam
Tinbergen Institute Amsterdam

June 2018

ABSTRACT

This paper investigates how biases in macroeconomic forecasts are associated with economic surprises and market responses across asset classes around US data announcements. We find that the skewness of the distribution of economic forecasts is a strong predictor of economic surprises, suggesting that forecasters behave strategically (rational bias) and possess private information. Our results also show that consensus forecasts of US macroeconomic releases embed anchoring. Under these conditions, both economic surprises and the returns of assets that are sensitive to macroeconomic conditions are predictable. Our findings indicate that local equities and bond markets are more predictable than foreign markets, currencies and commodities. Economic surprises are found to link to asset returns very distinctively through the stages of the economic cycle, whereas they strongly depend on economic releases being inflation- or growth-related. Yet, when forecasters fail to correctly forecast the direction of economic surprises, regret becomes a relevant cognitive bias to explain asset price responses. We find that the behavioral and rational biases encountered in US economic forecasting also exists in Continental Europe, the United Kingdom and Japan, albeit, to a lesser extent.

JEL classification: G14, F47, E44.

Keywords: Anchoring; rational bias; economic surprises, predictability; stocks; bonds; currencies; commodities; machine learning.

---

*We thank seminar participants at the Econometrics and Financial Data Science workshop at the Henley Business School in Reading, at the APG Asset Management Quant Roundtable, at the MAN-AHL Research Seminar in 2017 and at the Annual Conference of the Swiss Society for Financial Market Research (SGF) Conference in 2018 in Zurich for their helpful comments. We thank APG Asset Management and AHL Partners LLP for making available part of the data.

†Philip Stork (corresponding author) is at VU University Amsterdam, FEWEB, School of Finance and Risk Management, de Boelelaan 1105, 1081 HV Amsterdam, The Netherlands. Tel.: + 31 20 59 83651. E-mail address: p.a.stork@vu.nl.

1 Introduction

The presence of bias in analysts' forecasts is a widely investigated topic. Early literature focuses on the bias present in equity analysts' forecasting of earnings per share (FEPS), and attempts to explain why earnings estimates are systematically overoptimistic[1]. Only recently the same attention given to FEPS by the literature was given to the analysis of potential biases in macroeconomic forecasts. For instance, Laster et al. (1999) argue that forecasters have a dual goal: forecasting accuracy and publicity. Forecasters would depart from the consensus (which is typically accurate) when incentives related to their firms' publicity outpace the wages received by being accurate. The authors find this trade-off to vary by industry. Ottaviani and Sorensen (2006) compare two theories of professional forecasting, which lead to either forecasts that are excessively dispersed or forecasts that are biased towards the prior mean (herding)[2]. A drawback of this early literature on macroeconomic forecasts is that it fails to empirically test the direction and size of the bias, but mostly elucidates that dispersion of forecasts is plausible under different (sometimes stringent) assumptions. We also note that both previous papers focus on rational bias explanations for macroeconomic forecast rather than on cognitive issues.

To the best of our knowledge, Campbell and Sharpe (2009) are the first to address macroeconomic forecasts from both an empirical and behavioral bias approach. Their study hypothesizes that experts' consensus forecasts of economic releases are systematically biased towards the previous release. This bias is consistent with the adjustment heuristic proposed by Tversky and Kahneman (1974). This cognitive bias, commonly known as anchoring, is characterized by the human propensity to rely too heavily on the initial value (the "anchor") of an estimation when updating forecasts. In other words, individuals tend to make adjustments to original estimates that do not fully incorporate the newly available information. Thus, anchoring underweights new information in detriment of the "anchor".

In this paper, we investigate other potential biases embedded in macroeconomic consensus forecasts. The main hypothesis of our paper is that, beyond anchoring, other inefficiencies in the distribution of forecasts are informative in predicting economic surprises. More specifically, we hypothesize that some biases proposed by the literature are reflected in moments of the distribution of macroeconomic forecasts, such as the disagreement among forecasters (second moment) and skewness of forecasts (third moment). As market prices react to the information flow, economic surprise predictability might give rise to return predictability, as reported by Campbell and Sharpe (2009) and Cen et al. (2013). As a consequence, we conjecture that economic surprises as well as asset returns around these releases are predictable.

[1]For instance, De Bondt and Thaler (1990) suggest that equity analysts suffer from a cognitive failure which leads them to overreact and have too extreme expectations. At the same time, Mendenhall (1991) argues that underreaction to past quarterly earnings and stock returns contributes to an overoptimistic bias in earnings. Overreaction and underreaction as causes for an overoptimistic FEPS are, though, reconciled by Easterwood and Nutt (1999), who defend that analysts underreact to negative earnings announcements but overreact to positive ones.

[2]Ottaviani and Sorensen (2006) build on the reputational herding model of Scharfstein and Stein (1990), who suggest 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.

Our research builds on Campbell and Sharpe (2009), who hypothesize that surprises over economic releases are predictable, as they tend to underreact to new information. They find that the previous economic releases of 10 important US economic indicators explain up to 25 percent of the subsequent economic surprises[3]. Anchoring in forecasting seems not to be, however, restricted to macroeconomic data releases. Cen et al. (2013) show that anchoring also plays a significant role in FEPS of firms by stock analysts. Their study suggests that analysts tend to issue optimistic (pessimistic) forecasts when the firms' FEPS is lower (higher) than the industry median.

Further, Zhang (2006) investigates the link between anchoring, underreaction and information uncertainty. The author builds on the earlier post-earnings-announcement-drift (PEAD) literature (see, e.g., Stickel, 1991), which states that analysts underreact to new information when revising their forecasts due to behavioral biases, such as conservatism (Ward, 1982) or overconfidence (Kent et al., 1998). He suggests that a greater dispersion (disagreement) in analysts FEPS, which forms his proxy for information uncertainty, contributes to a large degree of analysts underreaction. Consequently, in an environment of high dispersion of FEPS, or for firms with greater information uncertainty, analysts will tend to incur larger positive (negative) forecasts errors and larger subsequent forecast revisions following good (bad) news[4].

Capistrán and Timmermann (2009) argue that the causality between underreaction and disagreement depicted by Zhang (2006) may also work the other way around. Capistrán and Timmermann (2009) argue that, as forecasters have asymmetric and differing loss functions, they react differently to macroeconomic news. In doing so, forecasters update their predictions in different ways and at different points in time as a reaction to the same news flow, giving rise to forecast disagreement. In line with Capistrán and Timmermann (2009), Mankiw et al. (2004) suggest that, as there are costs involved in gathering information and making adjustments to forecasts, experts underreact to recent news and only update their predictions periodically. Thus, in such a sticky-information model for forecasts adjustments, only part of the pool of forecasters would update their predictions at each period, also corroborating for dispersion of forecasts and information uncertainty. Interestingly, Zarnowitz and Lambros (1987) and Lahiri and Sheng (2010) use dispersion of forecasts as a measure of forecast uncertainty, not information uncertainty.

When attempting to find predictive value in disagreement measures among forecasters, Legerstee and Franses (2015) use the standard deviation of forecasts and the 5th and 95th percentile of survey forecasts to predict macroeconomic fundamentals. The 5th and 95th percentiles of survey forecasts (especially when used in combination with the mean or median) are, arguably, proxies for the skewness of forecasts, which is explicitly explored by Colacito et al. (2016) and

[3]Interestingly, it is yet unclear if this bias has a behavioral nature or if it is led by professional forecasters' strategic incentives.

[4]Note that another branch of the literature on equity analysts' forecasts proposes that biases are caused by strategic behavior, i.e., a rational bias, in line with Laster et al. (1999) and Ottaviani and Sorensen (2006) for macro releases. For instance, Michaely and Womack (1999) advocate that equity analysts employed by brokerage firms (underwriter analysts) often recommend companies that their employer has recently taken public. In the same vein, Tim (2001) suggests that a rational bias exists within corporate earnings forecasts because analysts trade-off this bias to improve management access (via positive forecasts) and forecast accuracy.

Truong et al. (2016). In their study, Colacito et al. (2016) use skewness of expected macroeconomic fundamentals to predict expected returns, whereas Truong et al. (2016) use the skewness of FEPS survey data to predict quarterly earnings.

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

Our contribution to the literature on forecasting bias is four-fold. First, we identify new biases in experts' expectations (over and above the anchoring bias), which are statistically significant predictors of economic surprises. More specifically, we are the first to empirically validate the rational bias hypothesis of Laster et al. (1999) and Ottaviani and Sorensen (2006) in a large multi-country data set of macroeconomic releases. Within such models, forecasters possess private information which is unveiled via the skewness of the distribution of forecasts. Second, by using a popularity measure per economic indicator and by expanding the number of countries and indicators tested vis-à-vis Campbell and Sharpe (2009), we show that the prevalence of biases is related to attention. This finding is supported by the fact that as we move from very popular economic releases, such as the Non-farm payrolls (NFP) employment number, Retail Sales and Consumer Confidence towards less watched indicators, biases become less pervasive. The same effect is observed when we compare our results for the US to those in other countries, in which economic indicators are forecasted by much fewer experts. Third, we confirm the hypothesis that, by predicting economic surprises, one can predict asset returns around macroeconomic announcements. We find that expected economic surprises can largely predict the direction of market responses around data releases in-sample and, to a lesser degree, out-of-sample. Hence, the expected component of surprises can help to explain market responses, thus building on the results of Campbell and Sharpe (2009), who suggest that markets respond to the unpredictable component of surprises. The explanatory power and predictability achieved by our models are higher for local equity and bond markets than for foreign markets, currencies and commodities, which is intuitive, as those markets are the ones more intrinsically linked to the fundamentals being revealed by macroeconomic indicators. The relation between asset responses and economic surprises is, though, distinct across the stages of the economic cycle and conditional on the nature of the information released, being either inflation- or growth-related. We find that on an out-of-sample basis, point-forecast is better performed by non-linear machine learning models as they seem to capture the dynamics of market responses around macroeconomic announcements better than linear regression models. Fourth, we are the first to recognize that a regret bias might influence how asset markets react to macroeconomic surprises.

The four key implications of our research are: 1) a better understanding of the "market consensus" and of the informational content of higher moments of the distribution of macroeconomic forecasts by regulators, policy makers and market participants; 2) the challenge of standard weighting schemes used in economic surprise indexes, which, we reckon, can be improved by changing from "popularity" (or "attention")-weighted to un-weighted; 3) the proposition that advanced statistical learning techniques should be used to refine the forecast of market responses amid macroeconomic releases and 4) the opening of a new stream in the literature to investigate regret effects in asset responses around announcements of forecasted figures.

The remainder of this paper is organized as follows. Section 2 provides a generic formulation of research applied to forecast biases. Section 3 describes the data and methodology employed in our study. Section 4 presents our empirical analysis and Section 5 concludes.

2 Forecast biases, anchoring and rationality tests

Let us first introduce the generic formulation of research applied to forecast biases, as used by Aggarwal et al. (1995), Schirm (2003) and Campbell and Sharpe (2009). In brief, this formulation consists of a rationality test in which macroeconomic forecasts are assessed to have properties of rational expectations. Such assessment is done, in its basic format, by running regressions with the actual release, A_t, as the explained variable, and the most recent forecast, F_t, as the explanatory variable, as follows:

$$A_t = \beta_1 F_t + \epsilon_t, \tag{1}$$

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

An alternative and more intuitive formulation of this rationality test, as suggested by Campbell and Sharpe (2009), can be achieved by subtracting the forecast from the left side of Eq. (1):

$$S_t \equiv A_t - F_t = \beta_2 F_t + \epsilon_t, \tag{2}$$

This manipulation yields to the forecast error or the "surprise", S_t, as the new explained variable, which is still dependent on forecast values. In Eq. (2), rationality holds when β_2 is not significantly different from zero; otherwise, a structural bias is perceived. For the specific case of anchoring, we can dissect the forecast bias using the following model:

$$F_t = \lambda E[A_t] + (1 - \lambda)A, \tag{3}$$

where $E[A_t]$ is the forecaster's unbiased prediction, and A is the anchor, which equals to the value of the previous release of the indicator of interest. In such a model, if $\lambda < 1$ so that $1 - \lambda > 0$, then the consensus forecast is anchored to the previous releases of the indicator. If $\lambda = 1$, no anchor is observed. By applying expectations to Eq. (2), then, substituting $E[A_t] = E[S_t] + F_t$ into Eq. (4a), we obtain Eq. (4d) after some manipulations:

$$F_t = \lambda(E[S_t] + F_t) + (1 - \lambda)A, \tag{4a}$$

$$\lambda E[S_t] = F_t - \lambda F_t - A + \lambda A, \tag{4b}$$

$$E[S_t] = \frac{F_t - \lambda F_t - A + \lambda A}{\lambda}, \tag{4c}$$

$$E[S_t] = \frac{(1 - \lambda)(F_t - A)}{\lambda}, \tag{4d}$$

assuming $\gamma = \frac{(1-\lambda)}{\lambda}$ and adding a intercept (α) we find[5]:

$$S_t = \alpha + \gamma(F_t - A) + \epsilon_t, \tag{5a}$$

$$S_t = \alpha + \gamma ESA_t + \epsilon_t, \tag{5b}$$

which reveals a direct test of anchoring, identified when the γ coefficient is positive, where ESA_t is the expected surprise given the presence of an anchor.

3 Data and Methodology

In this paper, we mostly employ ordinary least square (OLS) regression analysis adjusted for Newey-West standard error with the goal to offer interpretability to our results. More advanced statistical learning techniques are employed, but their usage is restricted to section 4.3.

We use macroeconomic release data from the ECO function in Bloomberg in our analysis. This data comprises of time-stamped real-time released figures for 43 distinct US macroeconomic indicators, as well as information on forecasters' expectations for each release. See Table 1 for an overview of these indicators. This expectations information comprises of 1) the previous economic release, 2) the cross-sectional standard deviation of forecasts, 3) the lagged median survey expectations, and the 4) the skewness in economists' forecasts, calculated as the mean minus median survey expectations. 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, thus covering 4,422 business days and 21,048 individual announcements. The consensus forecast is the forecast median, in line with Bloomberg's (and most other studies') definition.

[Please insert Table 1 about here]

[5]The above derivation builds fully on the work of Campbell and Sharpe (2009). The only difference between our approach and theirs lies on the fact that they consider the anchor to be the average value of the forecasted series over a number ($h \leqslant 3$) of previous releases, whereas our anchor variable relies only on the previous release (h=1). Robustness test for h>1 will be provided in future versions of this study.

We note that the economic indicators tracked are released in different frequencies and throughout the month. This a-synchronicity 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 states of the economy as a control variable in our empirical analysis, we also implement the Principal Component Analysis (PCA)[6]-based *nowcasting* method of Beber et al. (2015) using the same 43 distinct US macroeconomic indicators. Their *nowcasting* method allows us to access the real-time growth and inflation conditions present at the time of any economic release[7]. Table (1) provides details on stationary 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 in order to proxy for wealth effects and risk-appetite, respectively, as additional control variables in our empirical analysis.

3.1 Economic surprise predictive models

Following Eq. (5b), we hereby extend the *anchor-only* predictive model for economic surprises by incorporating moments of the distribution of macroeconomic forecasts and the control variables stated above. The moments of the distribution of macroeconomic 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). Eq. (6) is our *unrestricted economic surprise* model (*UnES* model):

$$S_t = \alpha + ESA_\varphi + SurvLag_\varphi + Std_\varphi + Skew_\varphi + Infl_\varphi + Growth_\varphi + Stocks_\varphi + VIX_\varphi + \epsilon_t, \quad (6)$$

where subscript φ (used hereafter) is *t-1*, *ESA* is the expected surprise given anchor[8], *SurvLag* is the lagged consensus forecast (the previous median of economic forecasts), *Std* is the dispersion (standard deviation) of economic estimates across forecasters, and *Skew* is the skewness of economic estimates across forecasters. *SurvLag*, *Std* and *Skew* are the three variables selected

[6]PCA is a unsupervised machine learning method that describes correlated variables into a set of orthogonal (linearly independent) variables, so-called principal components.

[7]The Beber et al. (2015) *nowcasting* method splits indicators among 4 categories (i.e., output, employment, sentiment, and inflation). We follow the same classification but we aggregate output, employment and sentiment indicator into a single category, i.e., growth. As our set of indicators perfectly matches the ones of Beber et al. (2015), this attribution exercise is straightforward. The only nuance that differs our *nowcasting* method from these authors' is that we use a single parameter to adjust for the non-stationarity of some series. Beber et al. (2015) adjust series using one-month and twelve-month changes, whereas we use six-month changes across all non-stationary indicators.

[8]The coefficient γ of Eqs. 5a and 5b is excluded from this model representation and subsequent ones for conciseness of presentation. We use the subscript φ (i.e, $t-1$) to clearly state that the model is predictive. In reality, the subscript t would still suggest a prediction as most macroeconomic indicator surveys close for forecast submission days before the economic release. For the case of Bloomberg, surveys close one business day prior to the data announcement.

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 consensus 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). *Skew* is used to test for the presence of strategic behavior and rational bias in macroeconomic 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.

3.2 Market response predictive models

Once predictive models of economic surprises are estimated via Eqs. (5b) and (6), we use the predictions to explain market responses between one minute before and one minute after ([t-1, t+1]) the release-time (t) of macroeconomic data. We do this by using the *expected* economic surprise produced by the different economic surprise models as an explanatory variable to forecast returns. We use three types of market response predictive models: 1) the *anchor-only* model, in which the expected surprise given anchor (*ESA*) is the only predictor of economic surprises, thus Eq. (6) with only one explanatory variable; 2) the *unrestricted* model, using all explanatory variables stated by Eq. (6); and 3) the unrestricted-extended response model, which entails the *unrestricted* model extended with a set of exogenous variables. The generic formulation of the two expected surprise-based models used is given by Eq. (7), whereas Eq. (8) specifies the 1) *anchor-only* model, as follows:

$$R_t = \omega + E(S_t) + \epsilon_t, \tag{7}$$

$$R_t = \omega + E(\alpha + \gamma ESA_t) + \epsilon_t, \tag{8}$$

where R_t is the market response calculated around the interval [t-1, t+1], thus the one minute before and one minute after the time the economic data is made available, and $E(S_t)$, the expected surprise, is derived from Eq. (5b).

The *unrestricted* response model is specified by Eq. (9):

$$R_t = \omega + E(\underbrace{\alpha + ESA_\varphi + SurvLag_\varphi + S_\varphi + Skew_\varphi + Infl_\varphi + Growth_\varphi + Stocks_\varphi + VIX_\varphi}_{\text{Unrestricted economic surprise model (UnES)}}) + \epsilon_t. \tag{9}$$

Eq. (10) provides a generic formulation of the *unrestricted-extended* response model because we do not implement it as an OLS regression only but also in the form of a Ridge regression and a Random forest model[9]:

$$R_t = \omega + E(UnES)_\varphi + ESA_\varphi + SurvLag_\varphi + Std_\varphi + Skew_\varphi + Infl_\varphi + Growth_\varphi + Stocks_\varphi + VIX_\varphi + \sum_{t=s}^{S} R_{t-s} + \epsilon_t, \tag{10}$$

where $UnES$ is the outcome of *unrestricted economic surprise model* of Eq. (6) and s=[5, 10, 20, 30, 40, 50, 60] minutes. For the Ridge regression model, we tune the shrinkage hyper-parameter (ϕ, typically called λ) via cross-validation using three splits of the train data set. For Random forest, we first run a cross-validation step for feature selection (using variable importance as guidance) and, then, tune the model by minimizing out-of-bag (OOB) errors[10] to obtain the parameter m for the number of random features considered at each branch split[11]. We allow the Random forest model to grow 500 trees per run.

We calculate the market responses across equity, treasury, currency, and commodity markets. More specifically, we use the following instruments: S&P500 index future, Euro-Stoxx index future, FTSE100 index future, 2-year US Treasury Note future, 2-year Bund future, 10-year Gilt future, Oil WTI future, Gold future, Copper future, GBPUSD forwards, JPYUSD forwards, CHFUSD forwards, AUDUSD forwards, EURUSD forwards, and CADUSD forwards. The market responses are calculated and used in our analysis for the entire history available per market instrument[12,13].

4 Empirical analysis and results

We split our empirical analysis and results section into five parts. Section 4.1 reports the results of predicting models for economic surprises. Section 4.2 dissects market responses as cumulative average returns (CAR) across multiple time-frames. Section 4.3 reports our findings of market response predictive models. Section 4.4 evaluates the presence of regret effects around macroeconomic announcements. Section 4.5 checks for the robustness of our findings.

[9]See Appendix A for details on Ridge and Random forest model.

[10]The usage of out-of-bag errors is an efficient replacement for cross-validation for tunning methods that rely on bootstrap to reduce the variance of a learning method. As such methods already make use of a bootstrapped subset of the observations to fit the model, whereas another subset of the observation is unused, the latter subset (so-called the *out-of-bag* (OOB) observations) can be used to calculate prediction error, thus, called OOB errors.

[11]Given the relative small number of observations available in our data set, we apply 100 repeats of our OOB-based tuning approach to obtain m, which is selected as the mode of the optimal m across all repeats.

[12]Response data is available since 18/9/2002 for the S&P500 index E-mini future, 22/6/1998 for the Euro-Stoxx index future, 1/1/1996 for the FTSE100 index future, 2/1/1996 for the 2-year US Treasury Note future, 10/5/1999 for the EUREX 2-year Bund future, 1/1/1996 for the 10-year Gilt future, 2/1/1996 for the NYMEX Oil WTI future, 2/1/1996 for the NYMEX Gold, 2/1/1996 for the NYMEX Copper, 1/1/1996 for GBPUSD forwards, 1/1/2000 for JPYUSD forwards, 1/1/2000 for CHFUSD forwards, 1/1/1996 for AUDUSD forwards, 16/7/1997 for EURUSD forwards, and 1/1/2000 for CADUSD forwards. This response data is provided by AHL Partners LLP.

[13]Return series for futures use first and second contracts for all markets. In general, return series use first contracts, which are rolled into second contracts between 5 and 10 days prior to the last trading day of first contracts, following standard market practice. Return series for currencies are calculated using synthetic one-month forwards.

4.1 Predicting economic surprises

In this section we report our findings from Eqs. (5b) and (6), i.e., 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 distinct US macroeconomic indicators analyzed. 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 ten percent 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.

Table (2) suggests that 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 percent of the times in our sample. This result is confirmed by the *unrestricted* model, in which ESA is statistically significant 67 percent of the times. The results for the *unrestricted* model reveal that the *Skew* factor is also often significant (72 percent) across our individual indicator regressions. This supports our conjecture that forecasters may behave strategically (a rational bias), which is in line with Laster et al. (1999) and Ottaviani and Sorensen (2006)[14]. *SurvLag* and *Std* are somewhat statistically significant, with 40 and 35 percent of the times, respectively. The result for *SurvLag* challenges our hypothesis that an anchor towards the previous consensus forecast holds empirically. The weak statistical significance of *Std* among our individual regressions also suggests that disagreement among forecasters and information uncertainty are linked to economic surprises. The control variables $Infl$, $Growth$, $Stocks$, and IV are significant between 7 and 33 percent of times, suggesting a somewhat weak relation between them and economic surprises.

[Please insert Table 2 about here]

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 percent for the *anchor-only* model and 17 percent for the *unrestricted* model (R^2 medians are 2 and 14 percent, respectively).

We report for the *anchor-only* regressions positive coefficients for the ESA factor in 81 percent of the times. The *unrestricted* model delivers a positively signed ESA coefficient in 88 percent of the times. Both results suggest a robust relationship between economic surprises and the anchor factor. The frequency of positive coefficients found for *Skew* is, however, even higher than for ESA. The *Skew* regressors are positive 93 percent of times across all

[14]We also apply Eqs. (5b) and (6) 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 check is to identify whether or not the median forecast is an inefficient predictor of economic surprise (versus the mean) and to test whether the predictive power of our *Skew* measure (i.e., mean minus median forecasts) vanishes through the use of the mean forecast within ESA. Our findings indicate that replacing the median by the mean in ESA changes our results marginally, being thus qualitatively the same as our main results.

regressions. *SurvLag* and *Std* are with 56 percent also largely positive but to a lesser extent than *Bias* and *Skew*. Our control variables are to an even lesser extent positive (between 23 and 51 percent). 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 are, thus, supportive of our hypothesis that a rational bias may be embedded in macroeconomic forecasting due to strategic behaviour of forecasters, which 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 is higher than 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 ∞, as the average R^2 across the *anchor-only* model is 3.7 percent, whereas for the *unrestricted* model it is 17 percent.

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

[Please insert Table 3 about here]

When the *unrestricted* model is used, *US Personal Income MoM SA* (45 percent) is the indicator with the highest R^2, followed by *US GDP Price Index QoQ SAAR* (41 percent), and *Adjusted Retail Food Service Sales* (36 percent). Most R^2 reach a double-digit level, in contrast with the *anchor-only* model. Most anchor coefficients are also statistically significant, in line with the *anchor-only* model. In line with earlier results, the *Skew* coefficients are mostly positive and statistically significant, whereas the coefficient sign is more unstable for the *SurvLag* and *Std* coefficients. The control variables within the *unrestricted* model are mostly statistically not significant, especially when inflation surprises are being forecasted.

More importantly, by analyzing individual models' results, we are able to explore an additional aspect of macroeconomic 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 a *Popularity weight* measure, which uses the sum of our popularity measure across all indicators as denominator. We also aggregate statistics in Table (3) using the nine most popular US economic indicator as employed by Campbell and Sharpe (2009)[15]. Overall, we find that the model quality is higher for popular indicators. The R^2 (AIC) weighted using our popularity measure for the *anchor-only* model is 4.0 percent (110), whereas the (unweighted) average R^2 (AIC) is 3.7 percent (923). For the *unrestricted*

[15]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 macroeconomic indicator this item is not part of our set of nine most popular US indicators.

model, the weighted R^2 (weighted AIC) is 17 percent (102), whereas the average R^2 (average AIC) is 17 percent (896). Hence, popular indicators seem better explained by our explanatory variables. If we compare the percentage of positive and significant coefficients across all models (see last two rows of Table (3)) with the same measure weighted by popularity and using the most popular indicator only, we observe that ESA and $Skew$ are more likely to hold with the correct sign among popular indicators. This result applies to both *anchor-only* model and the unrestricted model concerning ESA. Hence, we conjecture that the rational and behavioral biases modelled by ESA and $Skew$ are more present among popular indicators. This finding makes explicit that the bias in analysis here links to the active behavior of forecasters, not to their lack of action, as suggested by inattention-type of (behavioral) explanations advocated by Mendenhall (1991), Stickel (1991), Campbell and Sharpe (2009) and Cen et al. (2013), such as the anchoring bias.

4.2 Market responses around macroeconomic announcements

In the following we evaluate how asset prices of four different asset classes (equities, treasuries, foreign exchange, and commodities) behave around macroeconomic announcements. Because we primarily investigate US macroeconomic releases, we target reactions in US local markets and the EURUSD, the main USD currency cross. Hence, we analyze responses on the following assets: S&P500 index future, 2-year US Treasury future, and EURUSD forwards. Note that bond returns are adjusted to have the opposite signal so to be consistent with the expected response to surprises for equity returns and currencies. The response time-frames used (in minutes) are -60, -50, -40, -30, -20, -10, -5, -1, 1, 2, 3, 4, 5, 10, 20, 30, 40, 50, 60. Negative time-frames imply a time before the relevant economic release, whereas positive ones mean the minutes after the economic release.

We assess market responses around macroeconomic announcements by calculating cumulative average returns (CARs) and classifying responses around announcements as *good* or *bad* news to the asset. This way, we calculate CAR separately for announcements that had a *positive* or *negative* effect on the specific asset price.

Figure (1) illustrates market responses, separately for the S&P500 futures, the 2-year US Treasury futures, and EURUSD forwards in rows, whereas the first column displays plots of reactions to *good* news and the second column offers plot of reactions to *bad* news. The CARs around macroeconomic announcements for positive and negative responses across multiple time intervals are provided in Table (4), given in basis points (bps) and as a percentage of the CAR observed during the the one-hour before until one-minute after the macroeconomic announcement interval [t-60min, t+1min].

[Please insert Figure 1 about here]

For the S&P500 futures (Figure (1)) in row one, first column of Table (4), we observe that the largest part of the *positive* response happens around the macroeconomic announcement (which occurs between time-frames -1 and 1). The CARs from one-hour before the announcement until one minute after the news is roughly +/- 10 bps for positive and negative responses, respectively, whereas the response around the announcement is also approximately +/- 10bps. In fact, the average CAR observed around the announcement makes up for more than 100 percent of the overall CAR observed in the [t-60min, t+1min] interval (i.e., 108 for positive and 102 percent for negative responses). We conclude that pre-announcement drifts are, on average, of an opposite direction to the overall CARs observed. However, we note that the pre-announcement drifts are very small relative to the response observed within the [t-1min, t+1min] interval. Our results also suggest that one-minute after that releases are made public up to 60 minutes afterwards, there are only small post-announcement drift effects within the S&P500 futures as only 1 and 5 percent of the CARs around the announcement is observed within the [t-60min, t+1min] interval for the *positive* and *negative* responses. In brief, the *positive* and *negative* market responses for the S&P500 index future midst macroeconomic announcements have a similar pattern: almost no drift prior to the announcement, a jump at the announcement, and roughly a flat post-drift effect up to one-hour after the announcement. This CAR pattern suggests that no exploitable market underreaction to US macroeconomic news releases seems to be present within the local equity market. At the same time, as there is no pervasive pre-announcement drift observed, no evidence of leakage or usage of private information by market participants is found.

The CARs observed in different time-frames for the Euro-Stoxx and FTSE100 index futures show patterns similar to the ones found for the S&P500 index future. The pre-announcement drifts have the opposite direction to the response found close to the macroeconomic news release, both for *positive* and *negative* responses. The post-drift we observe is in the same direction as the response, but is in both markets of higher magnitude than the one found for the S&P500 index future, ranging from 10 to 24 percent of the CARs found in the [t-60min, t+1min] interval. This result seems to suggest that both Euro-Stoxx and FTSE100 index futures are less efficient than the S&P500 index future.

[Please insert Table 4 about here]

For the 2-year US Treasury futures (Figure (1), we see in row two, first column of Table (4)) that the *positive* response is also very distinct around the announcement. The average CAR from one-hour before the announcement until 30-minutes before the announcement is flat. There is some evidence of a pre-drift in the direction of the response from 30-minutes before the announcement until one-minute before the announcement of roughly 10 percent of the CARs observed in the [t-60min, t+1min] interval. The CARs observed for both the *positive* and *negative* responses in the interval [t-60min, t+1min] is between absolute 2.4 and 2.7 bps. Differently from equity markets, there is some evidence of a post-announcement drift, with an additional 0.6 and -0.3 bps (20 and 13 percent of the CARs experienced in the [t-60min, t+1min]) move expected after *positive* and *negative* responses, respectively. When treasury markets for the UK and Germany are evaluated (see Table (4)), we observe similar patterns for the post-announcement drift and response in the [t-1min, t+1min] interval, but no consistent pre-announcement drift. We note that the post-announcement drift for *positive* responses are consistently larger than the ones for *negative* responses.

We see that for the EURUSD (Figure (1), row three, first column of Table (4)), the responses are once again very distinct and concentrated closely around the macroeconomic announcement time. The CAR from one-hour before the announcement until one minute before the announcement is roughly zero. The CAR within the interval -1 minute and +1 minute is relatively large, roughly 5 bps, concentrating between 102 and 105 percent of the CAR observed in the [t-60min, t+1min] interval. Differently than observed for the treasury and equity markets, the post-announcement drift tends to be in the opposite direction of the response observed around the data releases, dampening between 5 and 21 percent of the CAR observed in the [t-60min, t+1min] interval. For other currencies, the pre-announcement drifts are on average small and inconsistent with each other and with the responses observed in the [t-1min, t+1min] interval. The post-announcement drifts are mostly in the opposite direction to the response around announcements for the *negative* responses. However, for *positive* responses, the post-announcement drift responses are in the same direction as the responses around the announcements for most currencies and only in opposite direction for the EURUSD and CADUSD.

We assess the CAR around US macroeconomic announcements for three commodities: WTI oil, gold, and copper. The most notable difference between our results for these commodities versus the other asset classes investigated is that the responses observed in the [t-1min, t+1min] interval for commodities concentrate less of the overall CAR observed in the [t-60min, t+1min] interval than for the previous three asset classes equity. The responses observed in the [t-1min, t+1min] interval range from 78 percent to 110 percent. Evidence of any pre- or post-announcement drift is very inconsistent across commodities and across *positive* and *negative* responses. The reason for such inconsistency might be that commodities are less clearly linked to the business cycle of a particular country compared to equities, treasuries and currencies. For instance, as countries may be consumers or suppliers of specific commodities, it is unclear how the macroeconomic announcements in a specific country, should affect the price of commodities[16].

Finally, we notice some common features observed from the Figure (1 for S&P500 futures, 2-year US Treasury futures and EURUSD. Firstly, when market responses are one standard deviation higher than the average reaction, markets mean-revert strongly by the following two minutes after the surprise and continue to do so for the following three minutes, though, less aggressively. Secondly, when market responses are one standard deviation lower than the average reaction, a post-drift in the subsequent two minutes after the surprise is observed. Further, volatility tends to increase prior to announcements for the S&P500 and 2-year US Treasury futures markets but not for the EURUSD market. Such increase in volatility starts even more than 30 minutes before announcements in the S&P500 futures markets, whereas for 2-year US Treasury futures, it happens only in the last 20 minutes before announcements.

[16] For stocks, it is also unclear how positive news impacts prices. Late in a tightening cycle (high inflation), good news is bad for equities, whereas at an early stage in tightening cycle (low inflation), positive macroeconomic news is definitely good for equities.

4.3 Predicting market responses

In this section we analyze the estimates from Eqs. (5a), (9) and (10), i.e., the *anchor-only* (*restricted*) model, the *unrestricted* model (used to forecast economic surprises), and the *unrestricted-extended* model. Table (5) reports R^2, *AIC*, the frequency of the expected surprise coefficients that are positive for the three OLS-based models employed, and hit-ratios as well as root mean squared error (RMSE) for all models. Hit-ratios and RMSEs are reported for our train and out-of-sample or test data set[17,18], whereas other statistics are calculated in-sample, i.e, using the full data set.

Table (5) reports that R^2 monotonically increases across the three OLS regression models as we move from the *restricted* model to more comprehensive models[19]. The magnitude of gains in R^2 across the three types of models suggests that the *unrestricted-extended* models have much higher explanatory power. On average, *anchor-only* models deliver R^2 of 1.5 percent, whereas *unrestricted* response models have R^2 of 2 percent on average. In contrast, *unrestricted-extended* models, which no longer are univariate models, post average R^2 of 29 percent.

The AIC statistic estimated across the different models challenges somewhat the results provided by R^2: complex *unrestricted* models are deemed less informative once penalties for complexity are applied. The AICs for the *restricted* model are 74 percent of the times lower than for the *unrestricted* model (indicating dominance of the *anchor-only* model), whereas the AICs for the *restricted* models dominate the AICs from *unrestricted-extended* models at all times. The same AIC dominance holds for the *restricted* models over the *unrestricted* models. Average AICs across these three types of models confirm these findings. AICs of Ridge models are, however, superior than the ones of their OLS counterparts, the *unrestricted-extended* models, indicating that model quality is improved by shrinkage. These first results indicate that in-sample fit is superior for the most complex models versus simpler models from an explanatory power perspective, but not from a parsimony perspective. Despite that, differences in AICs are not large, indicating that the superiority of small models on this criteria is not absolute.

[Please insert Table 5 about here]

When evaluating the coefficient signs of the expected surprise factor in market response predictive models[20], at first glance, we find that coefficients are mostly positive. Among *anchor-only* models, on average 57 percent of the coefficients for expected surprise are positive, whereas for *unrestricted* models this is 64 percent. Within larger models, such as *unrestricted-extended* ones, the percentage of positive coefficients for the expected surprises falls to 54 percent.

[17] The in-sample period extends through our full data set (i.e, from 1997 to 2016), whereas our out-of-sample period (our test data set) comprises of the latest 25 percent of observations of the full data set. The training data used for tuning (typically via cross-validation) of machine learning methods and estimation of models employed for out-of-sample forecasting uses the earliest 75 percent of observations of the full data set.

[18] Not all statistics are provided for the Ridge and Random forest model as they are not available or are not straight forward to estimate or aggregate.

[19] Note that both *restricted* and *unrestricted* models are univariate models.

[20] We expect expected surprise coefficients to be, in general, positive, as we expect that equities, commodities, currencies would typically appreciate in response of positive economic surprise. Bond returns are adjusted to have the opposite in order to be consistent with the other asset classes.

Further, we evaluate results coming from our OLS models by making a split between local (US) markets and foreign markets. We find that, from a R^2 perspective, the *unrestricted* and *unrestricted-extended* models of local markets seem to outperform foreign markets. The percentage of positive coefficients for the expected surprise variable is equal or higher for local markets than for foreign markets across all models. This is an intuitive results as we assume that local fundamentals should explain local markets more than local conditions explaining foreign markets. Though, for the US, due to its dominant economic position, this assumption might be weaker than for other countries.

When we assess model goodness of fit across the different asset classes evaluated, we find that R^2 for our three OLS models are much higher for the stock and bond markets. Within stocks, the Euro-Stoxx-based models are the ones with higher R^2. In bonds, the 2y Bund models are the one with highest explanatory power. Copper-based models have the highest R^2 in Commodities. Results from AIC and from currencies are more mixed. In univariate models, the percentage of positive coefficient for expected surprises are higher for stocks and bonds (always above 59 percent) than for other asset classes, in line with our expectations.

Further, R^2 is almost the same for growth and inflation indicators. Nevertheless, AIC points for a clear superiority of models' fit of growth-based indicators over inflation-based. The percentage of positive coefficients for the expected surprise variable is also consistently higher for growth indicator (between 57-69 percent on average), as it is always lower than 50 percent for inflation indicators. We think that this result is caused by positive growth surprises being less directly linked to subsequent increases in interest rates by central bank than positive inflation surprises, as higher interest rates typically produces negative shocks to equities and, more indirectly, commodities.

Weighing model fit outcomes using our popularity measure does not lead to additional insights as R^2 and AIC are nearly the same across models that rely on less-popular indicators and models that rely on popular indicators. The percentage of positive coefficients for $UnES$ is, though, clearly higher for popular indicators.

As we assess the performance of predictions made by the various models, our first impression is that *anchor-only* and *unrestricted* models does not convincingly beat a 50 percent hit-ratio out-of-sample, despite delivering a roughly 54 percent hit-ratio in the train data set. Out-of-sample hit-ratios are only slightly better than a coin flip for the *unrestricted-extended* (51 percent) among all models we use. The same applies to the average Ridge and Random forest models as they also post out-of-sample hit-ratios of around 51 percent. Interestingly, train hit-ratios for the *unrestricted-extended* seem to more heavily overstate out-of-sample hit-ratios than done by the train hit-ratios of the Ridge and Random forest models[21]. For instance, the average train hit-ratio for the *unrestricted-extended* model is 63 percent, whereas for the Ridge is 53 percent and 52 percent for Random forest. Random forest is the model that seem to overstate testing hit-ratios by train ones the least.

Across all *unrestricted-extended* frameworks, hit-ratios seem to be consistently higher for

[21]This result, which suggests overfitting by the OLS models, is what motivates us to apply shrinkage, as done by the Ridge regression.

models that forecast stocks returns versus models that predict other asset classes, especially currencies and commodities, matching our findings from R^2. Hit-ratios also suggest that models based on growth indicators do a better job at forecasting market direction than models based on inflation indicators. Further, train and test hit-ratio of models that use popular macroeconomic indicators are not consistently higher than hit-ratios for the average model.

Concerning RMSE, a first noticeable observation is that OLS *unrestricted-extended* models outperform all other models (including machine learning-based models) in train set but deliver higher average test RMSE than all other models. This is the case across popular and unpopular indicators and might be a symptom of overfitting. Machine learning-based models, however, report average train RMSEs that are higher than for the *unrestricted-extended* model but deliver much lower out-of-sample RMSEs, respectively $1.14\text{x}10^{-3}$ and $1.07\text{x}10^{-3}$ for the Ridge and Random forest model (versus $1.40\text{x}10^{-3}$ for the *unrestricted-extended* model). RMSEs are often higher for the in-sample period than for the out-of-sample period, which may be cause by the different level of markets' volatilities in the two sample splits. RMSEs also vary substantially across asset classes, which is explained by the adverse levels of return volatility of the asset classes used. As expected, RMSEs are the lowest for bonds and currencies and higher for stocks and commodities. Further, both train and test RMSEs are the lowest for models that use inflation indicators and predict local markets.

Our results also indicate that market responses created by announcements of popular macroeconomic indicators are less predictable than responses of unpopular indicators as train and test RMSEs are consistently higher for popular indicators, across all models. This finding suggests that, even if economic surprises in popular indicators are easier to forecast (as biases are more pervasive), their market responses are less anticipated (in RMSE terms) by the predictable part of economic surprises. As earlier reported, hit-ratios estimated do not suggest that popular indicators are more predictable either. Hence, to some extent, market participants seem to either discount the biases incurred by forecasters when trading around economic surprises of popular indicators or to have better models to predict surprises. These results are somewhat connected to Campbell and Sharpe (2009), who conclude that market participants "look through" forecasters' biases within ten popular US macroeconomic indicators[22,23,24].

When we dig into the drivers of forecasts produced by the *unrestricted-extended*, Ridge regression and Random forest models, we find that the $E(UnES)$ variable is important but only after a couple of market-based information, such as the 5-minute asset return prior to the announcement as well as the prior day level of the VIX index and stock returns. We base this conclusion in two metrics: the percentage of significant coefficients estimated by our

[22] We use nine out of the ten US macroeconomic indicators evaluated by Campbell and Sharpe (2009). The only indicator used by these authors and not by us is the New Home Sales statistic, as we do not include housing market data in our analysis.

[23] The average weight used to calculate popularity-weighted statistics is 2.3 percent, whereas the average weight of the indicator used by Campbell and Sharpe (2009) within such weighting scheme is 3.5 percent, denoting the use of very popular indicators by the authors.

[24] Qualitatively similar results are obtained when we perform the supervised learning approaches specified in section 4.3 as a classification problem rather than in a regression setting. These results are available under request.

unrestricted-extended OLS models and the *Importance* measure extracted from our Random forest models. In addition to these two metrics, we somewhat rely on the the percentage of positive coefficients from regression models to evaluate if the relation found between responses and the $UnES$, ESA, $Skew$ variables have the expected coefficient sign[25].

[Please insert Figure 6 about here]

Table (6) reports the percentage of positive coefficients for predictive models using the OLS and Ridge approaches (in the train and test data sets, respectively reported in Panels A and B). We observe that the estimated coefficients for $UnES$ and ESA are more often positive than negative, in line with our expectations. For $Skew$ this results is less strong, as within the OLS model this variable is only positive between 45 and 48 percent of times. Nevertheless, because, among all regressors, $Skew$ and $UnES$ are the most correlated variables (reaching a correlation of 0.8 for some of our macroeconomic indicators), the fact that $Skew$ is mostly negative might be simply the manifestation of multicollinearity in the regression model. We also find that $Stocks$ and R_{t-5} to be consistently positive, suggesting a positive serial correlation between returns during announcements and prior asset returns. In the case of $Stocks$, such relation might be linked to time-series momentum, which is typically captured in daily frequency data[26]. In the case of R_{t-5}, a positive coefficient indicates the presence of pre-announcement price drift in the direction of the economic surprise-led responses just few minutes prior to the data release, indicating potential leakage of information or short-term trading activity by informed investors.

Turning into the percentage of significant coefficients estimated by the *unrestricted-extended* OLS model, we find that $Stocks$ and R_{t-5} are the variables most strongly connected to asset responses amid macroeconomic announcements. Returns at other times frames (minutes) before announcements are also connected to returns during announcements, despite the fact that the direction of the relationship is not clear. Among non-market data based regressors, $UnES$ and Std are linked to market responses the most, indicating the relevance of $UnES$ for market predictions. Using the Random forest *Importance* measure as guidance (i.e., node impurity)[27], we find that $Stocks$ and R_{t-5} but also the VIX are highly relevant for predictions (see Table (6) and Figure 2). As reported by *Importance*, $UnES$ is the most relevant non-market data predictor used by Random forest. The fact that past returns have predictable power in forecasting returns around data announcements also adds to the pool of evidence in the literature of failure of the Efficient Market Hypothesis (EMH) on its weak form.

[Please insert Figure 2 about here]

In brief, we show that cross asset returns around US macroeconomic data announcement can

[25] We evaluate the percentage of positive sign for these three variables only as we do not have a prior for the relation between past returns and market responses around macroeconomic announcements. The same applies for the relation between return volatility and market responses around announcements.

[26] If found that positive (negative) responses amid positive (negative) data surprises might be strengthened by the existing positive (negative) time-series momentum, one could hypothesize that serial correlation in prices (i.e., momentum) is intensified by economic surprises in the same direction or a series of such surprises, i.e., serial correlation in surprises.

[27] See Appendix A for details.

be largely explained by variables that represent biases in the behavioral of forecasters, such as $UnES$ and ESA, as well as market-based variables, such as $Stocks$ and R_{t-5}. Beyond that, we find that these variables also have some out-of-sample predictability power. Explanatory power and predictability[28] is higher for local stocks and bonds than for currencies and commodities, whereas local markets are better predicted than foreign markets. Goodness of fit measures (R^2 and AIC) indicate that larger models deliver much higher explanatory power but, taking parsimony into account, bigger models are only preferred when regularized. Contrary to our results on predictability of economic surprises, in which popular indicator are found to be more predictable, we find that market returns around announcements of popular macroeconomic indicators are less predictable than responses provoked by unpopular indicators.

Beyond that, our results suggest that the (regularized) machine learning methods applied are superior at avoiding overfitting in our data set than simpler models, such as the OLS regression. No model applied consistently outperforms other methods on producing superior RMSEs and hit-ratios, however, Random forest dominates other method on point forecast as it consistently delivers lower RMSEs. We hypothesize that this result might be driven by the fact that Random forest is the only non-linear method among the models tested. Finally, the variable *Importance* measures calculated seems to challenge the myth that Random forest is a "black-box" method as it allows somewhat for model interpretation.

4.4 Market responses, skewness of economic forecasts and regret

In the previous sections, we observed that the skewness of economic forecasts is strongly and positively linked to economic surprises and market responses. In the following, we hypothesize that the relation between skewness of economic forecasts and market responses depends on failures of our skewness-based model in forecasting surprises. The rationale behind this hypothesis is that if market participants use experts' forecasts to trade, market responses might be adversely affected by the skewness of forecasts when they fail to correctly predict surprises. More specifically, we hypothesize that: 1) if a forecasted surprise fails to predict the direction of the realized surprise, then the correspondent market response is relatively large and in the opposite direction to the forecasted surprise (i.e., in line with the realized forecast); 2) if a forecasted surprise is in line with the realized surprise, then the subsequent market response is relatively small and in line with both the forecasted and realized surprise.

The intuition of Hypothesis (1) is that a regret effect takes places in asset markets in line with the models of Loomes and Sugden (1982) and Bell (1982). Therefore, investors that would be positioned in line with the expected surprises close their losses quickly after the economic release is made public. Concurrently, when realized surprises are in line with expected ones (Hypothesis (2)), no additional trading activity is expected from market participants that are holding such expectations, as it is likely they have positioned themselves according to their

[28] We consider hit-ratio as our predictability measure as RMSE cannot be adequately used to compare return forecast of assets with very distinct volatility as in our exercise. We used the median hit-ratio among *unrestricted-extended* models to rank the predictability across the different asset classes studied.

expectations ahead of the specific release[29]. One strong assumption made in this exercise is that the direction of the expected surprise is driven by the direction of the skewness of forecasts, which is in line with our estimated economic surprise models but not a result found for every single macroeconomic indicator in our analysis[30].

In order to test the Hypotheses (1) and (2), we specify the following regression models:

$$R_t = \alpha + Skew_t^{+} * S_t^{-} + \epsilon_t, \tag{11a}$$

$$R_t = \alpha + Skew_t^{+} * S_t^{+} + \epsilon_t, \tag{11b}$$

$$R_t = \alpha + Skew_t^{-} * S_t^{-} + \epsilon_t, \tag{11c}$$

$$R_t = \alpha + Skew_t^{-} * S_t^{+} + \epsilon_t, \tag{11d}$$

where R_t is the market response. $Skew_t^{+}$ is the skewness in forecasts when it is positive and $Skew_t^{-}$ when it is negative. S_t^{+} is the realized surprise when it is positive and S_t^{-} when it is negative. Hence, explanatory variables in these models are interaction between surprises and skewness in forecasts. To make the interpretation of the estimated coefficients of these interaction terms easier, we run regressions with the absolute value of these explanatory variables. Importantly, given our assumption that the direction of the expected surprise is driven by the direction of the skewness of forecasts, we interpret the variables $Skew_t^{+} * S_t^{-}$ and $Skew_t^{-} * S_t^{+}$ as scenarios in which the expected surprise failed to forecast the realized economic surprise. In the same line, $Skew_t^{+} * S_t^{+}$ and $Skew_t^{-} * S_t^{-}$ are scenarios in which the expected surprise was successful in forecasting the economic surprise.

Hence, these regressions split the direction of the skewness and realized surprises to map the four possible scenarios in which the responses can be evaluated: 1) the presence of positive skewness and negative economic surprise (skewness fails to forecast the direction of surprise); 2) the presence of positive skewness and positive economic surprise (skewness successfully forecasts the direction of surprise); 3) the presence of negative skewness and negative economic surprise (skewness successfully forecasts the direction of surprise); and 4) the presence of negative skewness and positive economic surprise (skewness fails to forecast the direction of surprise). The scenarios that give rise to regret are the ones in which the skewness fails to forecast the direction of economic surprise, thus, the scenarios number 1 and 4.

We note that Eqs. (11a) to (11d) do perform this four-scenario mapping by implementing each one of them as an individual univariate model. In order to have enough observations to run regressions for each of these four scenarios, we do not run regressions at the individual

[29] An implicit assumption embedded in Hypotheses (1) and (2) is that market participants that take part in an economic forecast survey also trade in asset markets (in line with their own forecasts) and that their forecasts influences market participants who trade in these markets.

[30] As indicated by Table (2), 93 percent of the estimated *unrestricted* models for economic surprises have a positive *Skew* coefficient.

economic indicator level but we aggregate observations for all economic releases. Aggregating surprise data for all the economic indicators within our sample is possible because surprises are also available as a number of standard deviations from the mean (apart from in *raw surprise* format)[31].

[Please insert Table 7 about here]

Table (7) reports the regression results for Eqs. (11a) to (11d) in an univariate setting. We observe that when economic surprises are negative, the percentage of coefficients found to be positive is lower than when economic surprises are positive. This indicates that negative surprises are more linked to negative market responses relative to positive surprises, when the absolute value of $Skew_t^+ * S_t^-$ and $Skew_t^- * S_t^-$ are used as regressors. This is in line with what we have expected[32]. However, when the skewness of forecasts is positive, the percentage of positive coefficients is the lowest (7 percent). This result suggests that negative responses are more frequently linked to negative surprises when the skewness of forecasts fails to correctly predict the economic surprise. This result is confirmed when we only use statistically significant coefficients in our analysis, as the percentage of positive and statistically significant coefficients for the $Skew_t^+ * S_t^-$ is zero versus 100 percent for $Skew_t^- * S_t^-$. We interpret this finding as being supportive of our hypothesis that regret affects market participants on trading around economic surprises.

In line with our expectations, when surprises are positive, the number of coefficients pointing towards a positive market response always exceeds 50 percent. Nevertheless, the number of coefficients pointing towards a positive market response is higher when the skewness of forecasts is negative (87 percent) than when it is positive (67 percent). This finding is confirmed when only statistically significant coefficients is taken into account in the analysis. This finding connects to our results when negative surprises are evaluated and supports our conjecture of a regret effect within economic surprises.

4.5 Market responses through the economic cycle

In section 4.3, we report that the coefficients for expected surprises from growth indicators as a predictor of market responses are more frequently positive than from inflation indicators, especially for equities. This observation is in line with our notion that inflation surprises are more likely to have an adverse impact on equities than economic surprises coming from growth indicators relative to bonds. In other words, we think that surprises from growth (inflation) indicators are generally connected to positive (negative) responses of equity markets, whereas surprises coming from either inflation and growth indicators likely have a negative impact on bonds. Nevertheless, we can also hypothesize that, in certain periods, inflation surprises are

[31]The application of these regressions using *raw surprise* would be biased as the different magnitude of the multiple economic releases would create a biased relation between the explanatory variable (surprises) and the explained variable (market response).

[32]We note that the coefficient sign of the bond returns is reversed by us to be consistent with equity returns and currencies to what the expected direction of returns given economic surprises is concerned.

more strongly connected to positive equity surprises, such as during deflationary periods. These assumptions lead us to conjecture that the impact of macroeconomic indicators in market prices is time-varying and likely linked to the different stages of the economic cycle. For instance, we would expect that positive inflation surprises negatively impact equity returns more strongly during later stages of the economic cycle (when inflation is already high) than during its early stages or recessionary periods. Hence, we hereby evaluate the influence of the business cycle into asset market responses amid economic surprises. As our main experiment refers to the US, we carry out this test for this country alone.

Prior to this test, we must characterize how we split the economic cycle into its different stages. We make this separation by characterizing the economic cycle as a combination of three phases: recessions, recoveries and expansions. We identify recessions by following the classification of the US National Bureau of Economic Research (NBER). Recoveries are, then, identified as non-recessionary periods for which the US core inflation (CPI) is below two percent, whereas expansions are categorized as non-recessionary periods for which core inflation is equal or above two percent.

Subsequently, we apply the following univariate regression model for market responses for the different asset classes and stages of the economic cycle separately:

$$R_{t,m,s} = \alpha + S_{t,m,s} + \epsilon_t, \tag{12}$$

where $R_{t,m,s}$ is the market response calculated around the interval [t-1, t+1], for markets m = 1...4 and stage of the economic cycle s = 1...3, whereas $S_{t,m,s}$ are realized economic surprises. Note that Eq. (12) is identical to Eq. (7) with the exception that economic surprises here are realizations, rather than expected one. As done in section 4.4, we run regressions using observations for all releases available to be able to apply them to a large enough sample. Running a regression per economic indicator after splitting our sample into the three phases of the economic cycle would likely lead us to poor statistical results.

We report the aggregated (average) coefficients for $S_{t,m,s}$ per asset classes and per stage of the cycle in the Panel A of Figure 4. Our first observation is the larger size of coefficients obtained for equities versus other asset classes, which links to the higher volatility of this asset class versus bonds and FX. We suspect that the tamed average coefficients for commodities and FX is caused by the constituents of these two asset classes to be uncorrelated to each other, hence, exhibiting an unclear pro- or anti-cyclical orientation[33]. Secondly, we observe that economic surprises are associated with larger equity responses during recessions than during recoveries and expansions. This finding suggests that amid periods of higher economic growth the impact of positive economic surprises into equities fades. Potential explanations for this relation are the fact that higher economic growth may have already fueled into higher inflation during late stages of the cycle or equities may have become expensive by then. Contrary to equities,

[33]For instance, note from Table 5 that some commodities and currency pairs, such as Gold, USDAUD and USDCAD have very distinct percentage of positive coefficients linking expected economic surprises and market responses relative to the other commodities and currency pairs used.

positive economic surprises are related to negative bond returns, as expected. The magnitude of bond responses across the different stages of the economic cycle are, though, similar[34]. For FX, average coefficients are all slightly negative but more so during recessions, which may reflect USD weakness amid positive surprises as it tends to strengthen during recessions. For commodities, economic surprises are positive and larger during recessions fading as the cycle enters into the recovery and expansion phases. During expansions, economic surprises are not associated with higher commodity prices on average. We think that the reason for these market response patterns for commodities is similar to the one driving reactions for equities: the marginal impact of economic surprises diminishes amid higher economic growth or stretched commodity valuations[35].

Additionally, because we also hypothesize that inflation surprises may have an adverse effect on some asset classes, such as equities, we run Eq. (12) using inflation indicators only. Our results are presented in Panel B of Figure 4. Our first and most striking observation is that positive inflation surprises are connected to equity returns in a very different way compared to how economic surprises are connected to equity returns in general. Specifically, positive inflation surprises are related to very negative equity market responses during recessions. Despite this adversity, this relation seems intuitive as positive economic surprises during recession can categorize stagflation, when equities clearly perform badly. The recovery phase is when equities returns are the least negatively linked to positive inflation surprises, which is intuitive too, as inflation levels are typically not a concern in this period. During the expansion phase, when inflationary pressures build, equities again perform poorly amid positive inflation surprises, which is also in line with what we would expect. Bonds are negatively linked to positive surprises across all phases of the cycle, but less so during expansions, which may be justified by marginal decreasing responses given higher levels of inflation in this phase. For FX, positive inflation surprises are positively connected to USD strength during recessions and roughly unconnected to market prices during recoveries and expansions, which might be a result of the different currency pairs used having very distinct exposure to the economic cycle, as mentioned. During recessions, though, FX seems to react less negatively when only inflation releases are used. Similarly to equities, commodities responses show very different relation with inflation surprises in comparison to when we use surprises from all economic indicators. Commodities are most negatively linked to positive inflation surprises during recessions, which again may be connected to a bleak prospectus of commodities during stagflations. During recoveries, positive inflation surprises remain negatively connected to commodity prices. Though, positive inflation surprises are connected to higher commodity prices during expansions, indicating that higher inflation during this phase may be linked to upward pressure in resources and suggesting that commodities may have inflation hedge properties.

[34] Opposite to our analysis in section 4.4, the returns for bonds here do not use the reverse sign so coefficients obtained are consistent with ones for equities.

[35] Our results are qualitatively the same when we carry out a similar exercise for local markets (equities and bonds) only and growth indicator-only. Note that growth indicators are 33 out of 43 used in our analysis.

4.6 Robustness tests

4.6.1 Economic surprise models across regions

As a robustness test, we apply Eqs. (5a) and (6) across other regions, namely, Continental Europe, the United Kingdom and Japan[36]. Table (8) indicates that our results for these three regions are qualitatively the same as the ones reported for the US: *unrestricted* models tend to improve the R^2 of *anchor-only* models and the coefficients for the ESA and $Skew$ factors are mostly positive (the expected sign). These two coefficients are positive between 56 and 75 percent of all times, which is however lower than the percentage of correct signs found for the US. Yet, among coefficients for all factors (including control variables), ESA and $skew$ remain the ones that are mostly positive. Moreover, in significance terms, 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 are, in general, lower than for the US.

[Please insert Table 8 about here]

We conjecture that the difference in presence of biases in macroeconomic forecasting across the different regions might be explained by the number of experts dedicated to macroeconomic forecasts across these countries. The average number of analysts providing forecasts across all indicators and through our sample is 44 for the US. For Europe, Japan and the United Kingdom this number is, respectively, 9, 13, and 15. 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) forecasters possess private information; 3) such private information is revealed by the skewness in forecasts, given strategic behavior by experts.

4.6.2 Economic surprise models through time

We check how the strong relations found between economic surprises and the ESA and $Skew$ factors in our main analysis behave through time. To perform this stability test we employ a panel regression version of the Eq. (6), our main predictive model for economic surprises. We use a panel regression[37] 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 have normalized them into z-scores.

Our results are reported in Figure 4, which depicts the coefficient values and p-values of predictors ESA, $Skew$, Std and $SurvLag$ through time. We observe that the coefficients for variables ESA and $Skew$ are positive and statistically significant, with few exceptions. At the same time, the signs of the coefficients for Std and $SurvLag$ are unstable, fluctuating between

[36] The overview of macro releases for these regions can be provided under request

[37] 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.

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 $Skew$, indicating a much larger impact of the normalized ESA into the estimated economic surprise. Nevertheless, the magnitude of the coefficient for $Skew$ steadily increases through our sample and after 2009 it becomes structurally larger than the one for ESA. This behavior reiterates our main findings, which suggest that both ESA and $Skew$ are strongly connected to economic surprises. It also indicates that the $Skew$ factor has gained relevance lately, whereas the impact of our anchor-based factor (ESA) has diminished through the sample evaluated.

[Please insert Figure 4 about here]

4.6.3 Expected and unexpected economic surprises

As an additional robustness test, we compare how expected economic surprises are linked to market responses vis-à-vis to unexpected surprises. As earlier mentioned, expected surprises are the ones that can be predicted, in line with Campbell and Sharpe (2009). In this study, we have used the *anchor-only* and *unrestricted* models, as given by Eqs. (5b) and (6), to estimate expected economic surprises. In contrast, unexpected surprises are the residual component of surprises. In other words, the unexpected surprise is the part of the economic surprise that cannot be forecasted. We link market return around macroeconomic data announcements with expected surprises and unexpected surprises via the following *two-component response* model:

$$R_t = \omega + \delta_1 E[S_t] + \delta_2 (S_t - E[S_t]) + \epsilon_t, \tag{13}$$

where $E[S_t]$ can be provided by either the *anchor-only* model or by the *unrestricted* economic surprise model, and S_t is the realized surprise. Essentially, the goal of running such a regression model is to understand and compare whether and how market prices react to the two-components of economic surprises: the predictable component of surprises (by evaluating δ_1), and the unpredictable portion of economic surprises (by evaluating δ_2.).

In order to compare how market responses can be explained by expected surprises versus unexpected surprises, we also run the following *unexpected-surprise* response model:

$$R_t = \omega + \delta_3 (S_t - E[S_t]) + \epsilon_t, \tag{14}$$

The estimates of Eqs. (13) and (14) are provided by Table (9) aggregated across the 43 US economic indicators used. Panel A reports the R^2 as well as the percentage of δ_1 and δ_2 coefficients that are significant for both the *anchor-only* (two-components) response model and the *unrestricted* (two-components) response model. Our findings suggest that the coefficient for unexpected surprises is often significant, on average 62 and 59 percent of all times, respectively for the *anchor-only* and *unrestricted* models. Nevertheless, the expected surprises are also frequently significant, but less so than for unexpected surprises. Still, because expected surprises are significant in 22 and 32 percent of all times, we conclude that market responses are also strongly linked to expected surprises, not only unexpected ones. The fact that expected surprises are more significant in *unrestricted* models versus the *anchor-only* model, and that unexpected surprises are less significant for the *unrestricted* models reiterate our results that factors beyond ESA are informative in predicting surprises, such as $Skew$.

[Please insert Table (9) about here]

Panel B of Table (9) compares the R^2 of univariate unexpected response models for the *anchor-only* and *unrestricted* response models with the explanatory power of their two-component (multivariate) counterparts in an aggregated basis. At first glance, we see that the average explanatory power delivered by the *anchor-only* and unrestricted-based unexpected surprise models are quite similar, suggesting that the unexplained portion of the surprise as modelled by these two approaches is comparable. When we compare the average R^2 delivered by the two-component models with the ones delivered by unexplained surprises only (across *anchor-only* and *unrestricted* models), it seems that R^2 for the two-component models is only marginally higher, by roughly one percent. Nevertheless, when we evaluate the percentage gain in R^2 delivered by the two-component models (versus the unexpected-surprise models), there is an indication that the two-component models increase the explanatory power of the unexpected-surprise models quite substantially. This gain is roughly 50 percent on average, across *anchor-only* and the *unrestricted* model. This finding indicates that the predictable portion of surprises adds substantial explanatory power to response models relative to the explanatory power of unexpected surprises-only models. These results contribute to our findings that expected surprise models comprise a relevant source of information on estimating market responses around economic surprises.

Further, for expected surprises the percentages of significant coefficients is roughly the same for the popularity-weighted and the un-weighted averages. This findings are in line with our earlier observation that responses connected to surprises of popular indicators are just as predictable than responses provoked by unpopular indicators. In contrast, we observe that the percentages of significant coefficients for unexpected surprises is higher for popularity-weighted and the most popular indicator than for the un-weighted average across *anchor-only* and *unrestricted* surprise models. In the same vein, the percentage gain in R^2 delivered by the two-component models is lower for popular indicators than for the average indicator, reflecting that the explanatory power added by expected surprises (on top of the R^2 produced by the unexplained portion) for popular indicators is less than for the average indicator. These findings indicate, despite expected surprises being connected to market response across indicators of any level of popularity, for popular indicators, the unexpected component of surprises is more prevalent than the expected one than for un-popular indicators. In other words, the unexpected component of surprises for popular indicator dominates their expected component, which is happens at a lower frequency for un-popular indicators.

In this way, our findings diverge from the bold conclusion of Campbell and Sharpe (2009) that traders "look-through" the bias. Nevertheless, our results differ from theirs partially because we use a much broader set of (un-popular) economic indicator. When we focus on the set of popular indicators used by these authors our results are more in line with theirs, in which unexpected component of surprises is the main explanatory variable of market responses (despite not being the only one). Hence, the link between expected surprises and market responses is more prevalent in a set of less popular indicators, despite the fact that biases are more pervasive on popular indicators.

5 Conclusion

This paper investigates how forecasters biases, both cognitive and rational, are associated with future macroeconomic surprises and their respective market responses around announcements in the US. We empirically confirm that the anchor bias previously recognized in the literature remains pervasive but we show that higher moments of the distribution of economic forecasts are also informative on predicting surprises. Particularly, our results suggest that the skewness of the distribution of economic forecasts provides reliable information to predict economic surprises. Whereas anchoring has clear behavioral roots, we assume that the information contained in the skewness of forecasts reflects a rational bias. This assumption builds on the literature on strategic behavior by forecasters, who have dual and contradicting objectives, i.e., forecast accuracy and publicity. According to this stream of research, forecasters typically stay close to the "pack" (herding) but eventually, when in the possession of what they perceive to be private information, they intentionally issue off-consensus forecasts, contributing to a highly skewed distribution of forecasts. Our results, thus, suggest that professional forecasters often possess private information and that they do make use of it by issuing controversial (and informative) forecasts. Additionally, we find that the relevance of the rational bias in predicting economic surprises seems to steadily increase through time and relative to the anchor bias. Under these conditions, macroeconomic surprises are, at least, partially predictable. Predictability is, though, stronger for popular indicators, suggesting that as we move from widely followed indicators towards less watched ones, biases become less pervasive. In the same vein, we show that predictability and the strong link between macroeconomic surprises and forecast skewness also holds for other countries/regions, such as Continental Europe, the United Kingdom and Japan, however, to a lesser extent.

A possible consequence of economic surprises being predictable is that responses observed in asset returns around macroeconomic announcements are also predictable. Our findings confirm this hypothesis in-sample and, to a lesser degree, out-of-sample. We identify that forecasts made using our *unrestricted-extended* economic surprise models outperform simpler models on forecasting market responses across the four asset classes studied. The success of these models is partially associated to the expected economic surprises modelled and partially linked to past returns, challenging the Efficient Market Hypothesis (EMH).

Asset returns around announcements of highly-followed macroeconomic indicators are as predictable as around releases of unpopular indicators, despite economic surprises being more predictable for popular indicators. Nevertheless, market responses around announcements of popular indicators are more frequently linked to the unexpected component of surprises than unpopular indicators are. We report that returns of assets that are sensitive to the fundamentals being revealed by macro announcements (local equities and bonds) are more predictable around such events than foreign markets, currencies and commodities. In addition, economic surprises are found to link to asset returns very distinctively through the stages of the economic cycle, and they strongly depend on whether economic releases are inflation- or growth-related. Our results also suggest that machine learning techniques outperform OLS regression models in point forecast, possibly due to their non-linear nature. Undoubtedly, the regularized machine learning models applied by us are superior than OLS regression at avoiding overfitting in our data set.

Yet, when forecasters fail to correctly forecast the direction of the economic surprises, another bias seems to play a role in explaining market responses: regret. We identify the presence of this cognitive bias as we find that negative (positive) market responses are more pervasive when the skewness of forecasts failed to correctly forecast surprises. Future research is warranted to strengthen our conclusions on the matter. Extending our findings to other cases, such as the forecasting of quarterly earnings releases, seems a natural next step to take. Improving our skewness-based forecast approach by the use the skewness of top-quartile forecasters also may strengthen our findings further.

We conclude with the four key implications of our findings: 1) a better understanding of the "market consensus" and of the informational content of higher moments of the distribution of macroeconomic forecasts by regulators, policy makers and market participants; 2) the challenge of standard weighting schemes used in economic surprise indexes, which we find can be improved by changing from "popularity" (or "attention")-weighted to un-weighted, as market responses around announcements of popular indicators are not more predictable than responses around releases of unpopular indicators; 3) the proposition that advanced statistical learning techniques should be used to refine the forecast of market responses amid macroeconomic releases, especially when such methods prevent overfitting and are somewhat transparent; and 4) the opening of a new stream in the literature to investigate regret effects in asset price responses around announcements of forecasted figures.

A Appendix

A.1 Ridge regression

The Ridge regression (Hoerl and Kennard, 1970) is a shrinkage method, similar to the Least Absolute Shrinkage and Selection Operator (Lasso) of Tibshirani (1996). The main difference between the Lasso and Ridge regression is that the former translates each coefficient by a constant factor ϕ (typically λ), truncating at zero, whereas the latter applies proportional shrinkage. The main consequence of such difference is that shrank coefficients by Lasso equal to zero, whereas for Ridge regression coefficient approach zero as limit. Therefore, the Ridge regression only applies shrinkage and not shrinkage and *variable selection* simultaneously, which should help forecast accuracy but does not improve model interpretation as the Lasso does. Thus, regression models shrank by Lasso are *sparse* version of the original regression model, whereas Ridge regressions are not.

The regression coefficients obtained by the Ridge methodology applied (β_{θ}^{L}) are estimated by minimizing the quantity:

$$\sum_{i=1}^{n}(y_1 - \beta_0 - \sum_{j=1}^{p}\beta_j x_{ij})^2 + \phi\sum_{j=1}^{p}\beta_j^2 = RSS + \phi\sum_{j=1}^{p}\beta_j^2 \tag{A.1}$$

where ϕ is the tuning parameter, which is estimated via cross-validation. The cross-validation applied uses three equal-size splits of our train data set. For a comparison between Ridge regressions and the Lasso, see Hastie et al. (2008).

A.2 Random Forest

Random forest is a decision tree-based method derived from bootstrap aggregation (i.e., bagging). Bagging entails fitting a regression many times by applying bootstrap to the train data set and, then, averaging the predictions of each model. The goal of applying bagging is to reduce a predictions' variance by averaging. As decision tree typically suffer from high variance, ensemble methods such as bagging, Random forest and boosting are warranted for better predictive accuracy. As mentioned, Random forest builds on bagging by growing a large collection of trees with the nuance that they are imposed to be as de-correlated as possible. Similarly to bagging, after de-correlated trees are grown as a 'random forest', then, predictions coming from them are averaged into a single prediction.

The Random forest (regression) predictor is given by:

$$\hat{f}_{rf}^{B}(x) = \frac{1}{B}\sum_{b=1}^{B}T(x;\Theta_b) \tag{A.2}$$

where, B is the number of $T(\cdot)$ trees grown, Θ_b is the set of characteristics of the bth tree to what it concerns the split variables, cut-points at each node and terminal-node values and x is the set of explanatory variables. For details on decision tree, which is the basic building block for Random forest, see Hastie et al. (2008).

A.3 Variable *Importance* measure

The *Importance* measure applied in our study[38] computes the average node impurity across all trees grown by the Random forest, reflecting how optimum partitions made by the different explanatory variables at each node compare with a 'pure' node, i.e., a constant fit over the entire region.

The natural starting point for the calculation of the *Importance* measure (Ψ_v) per variable $v = 1...V$ is a single decision tree T as follows:

$$\Psi_v^2(T) = \sum_{j=1}^{J-1} \hat{\iota}_j^2 I(w(j) = v) \tag{A.3}$$

where, the sum collects the importance of each variable v across the total number of nodes J of the tree. At each node j the input variable in analysis splits the region into two sub-regions associated with a boundary level, where sub-regions are either equal or higher than or smaller than the boundary level utilizing function $w(\cdot)$. The variable chosen to make this split is the one that maximizes the improvement (versus the previous nodes in this branch) $\hat{\iota}_j^2$ in squared errors achieved in this node relative to the decision of no partition, i.e., a 'pure' node.

As in Random forest many trees are utilized, the *Importance* measure per tree ($\Psi_v^2(T)$) has to be aggregated into a overall model *Importance* measure, which is achieve by averaging *Importance* across the total number of tree M:

$$\Psi_v^2 = \frac{1}{M} \sum_{m=1}^{M} \Psi_v^2(T_m). \tag{A.4}$$

For more details on *Importance* measures see Hastie et al. (2008).

[38]Note that there are alternative *Importance* measures that can be used in Random Forest and in other tree-based algorithms.

References

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

Beber, A., Brandt, M., Luisi, M., 2015. Distilling the macroeconomic news flow. *Journal of Financial Economics* 117 (3), 489–507.

Bell, D., 1982. Regret in decision making under uncertainty. *Operations research* 30 (5), 961–981.

Campbell, S., Sharpe, S., 2009. Anchoring bias in consensus forecasts and its effect on market prices. *Journal of Financial and Quantitative Analysis* 44 (2), 369–390.

Capistrán, C., Timmermann, A., 2009. Disagreement and biases in inflation expectations. *Journal of Money, Credit and Banking* 41 (2), 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 (1), 47–76.

Colacito, R., Ghysels, E., Meng, J., Siwasarit, W., 2016. Skewness in expected macro fundamentals and the predictability of equity returns: Evidence and theory. *Review of Financial Studies* 29 (8), 2069–2109.

De Bondt, W., Thaler, R., 1990. Do security analysts overreact? *American Economic Review* 80 (2), 52–57.

Easterwood, J., Nutt, S., 1999. Inefficiency in analysts' earnings forecasts: Systematic misreaction or systematic optimism? *Journal of Finance* 54 (5), 1777–1797.

Hastie, T., Tibshirani, R., Friedman, J., 2008. The elements of statistical learning: data mining, inference, and prediction (2nd ed.), Springer–Verlag.

Hoerl, A., Kennard, R., 1970. Ridge regression: biased estimation for nonorthogonal problems. *Technometrics* 12 (1), 55–67.

Kent, D., Hirshleifer, D., Subrahmanyam, A., 1998. Investor psychology and security market under- and overreactions. *Journal of Finance* 53 (6), 1839–1885.

Lahiri, K., Sheng, X., 2010. Measuring forecast uncertainty by disagreement: The missing link. *Journal of Applied Econometrics* 25 (4), 514–538.

Laster, D., Bennett, P., Geoum, I., 1999. Rational bias in macroeconomic forecasts. *Quarterly Journal of Economics* 114 (1), 293–318.

Legerstee, R., Franses, P., 2015. Does disagreement amongst forecasters have predictive value? *Journal of Forecasting* 34 (4), 290–302.

Loomes, G., Sugden, R., 1982. Regret theory: an alternative theory of rational choice under uncertainty. *Economic Journal* 92 (4), 805–924.

Mankiw, N., Reis, R., Wolfers, J., 2004. Disagreement about inflation expectations. *NBER Macroeconomics Annual 2003* 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 (4), 653–686.

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

Scharfstein, D., Stein, J., 1990. Herd behavior and investment. *American Economic Review* 80 (3), 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.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society* 58 (1), 267–288.

Tim, T., 2001. Rationality and analysts' forecast bias. *Journal of Finance* 61 (1), 369–385.

Truong, C., Shane, P., Zhao, Q., 2016. Information in the tails of the distribution of analysts' quarterly earnings forecasts. *Financial Analysts Journal* 73 (5), 84–99.

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

Ward, E., 1982. Conservatism in human information processing. In Daniel Kahneman, Paul Slovic and Amos Tversky. (1982). *Judgment under uncertainty: Heuristics and biases* New York: Cambridge University Pres.

Zarnowitz, V., Lambros, L., 1987. Consensus and uncertainty in economic prediction. *Journal of Political Economy* 95 (3), 591–621.

Zhang, X., 2006. Information uncertainty and analyst forecast behavior. *Contemporary Accounting Research* 23 (2), 565–590.

Table 1: Overview of US macro releases

This table reports the 43 US macroeconomic 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 macroeconomic indicator begins. Column *Frequency* reports in which frequency the indicator is released, where Q stands for quarterly, M stands for monthly and W stands 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 modelled using our methodology.

#	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	23:00:00 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
43	US Personal Cons. Exp. Price YOY SA	Inflation	05/02/10	M	14:30:00 GMT	1	Yes

Table 2: Aggregated results of anchor-only (restricted) and unrestricted economic surprise models for the US

Panel A reports the percentage of statistically significant coefficients across anchor-only and unrestricted regression models for economic surprises of US macroeconomic indicators. For example, 0.65 found for the ESA variable within the anchor-only model means that 65 percent of the ESA across the individual regressions run for the 43 US macroeconomic indicators are statistically significant at the 10 percent level. Panel B reports the percentage of positive coefficients across anchor-only and unrestricted regression models for economic surprises of US macroeconomic 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).

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

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

The table below 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 asterisks ***, **, and * indicate significance at the one, five, and ten percent level, respectively. 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.

Model	Anchor-only model				Unrestricted model													Popularity
Statistics / Regressors	R^2	AIC	Intercept	Anchor	R^2	AIC	Intercept	Anchor	Std	SurvLag	Skew	Inflation	Growth	Stocks	VIX	x R^2 gain	weight
US Initial Jobless Claims SA	0%	22901	128.8	-0.1**	7%	21925	-3586.0	-0.1	-0.2	0.0	2.1***	582.8	358	-24468	253***	∞	2.0%
US Employees on Nonfarm Payrol	0%	6151	-12**	-0.1	8%	5884	50**	0.0000	-0.001**	-0.0001	0.003***	-5	1	-172	-1*	∞	4.4%
U-3 US Unemployment Rate Total	0%	-2451	0.0***	0.1	11%	-2362	0.0	0.0	0.1	0.0**	1.2***	0.0	0.0	0.0	0.0	∞	4.3%
US Employees on Nonfarm Payrol	2%	5003	-5009***	-0.2*	16%	4937	1235	0	-1***	0**	2***	-592	1971**	122767	99	8	0.9%
US Continuing Jobless Claims S	2%	17944	2	0.3***	8%	17910	30***	0***	0	0***	0	0	-2**	-106	1***	4	0.3%
ADP National Employment Report	1%	3132	3786.5	0.1	13%	3129	49**	0	0	0	0**	4	-1	506	-1	13	0.9%
US Average Weekly Hours All Em	0%	-183	0.0	-0.1	27%	-196	0.0	0.4*	0.8*	0.0	2.3***	0.0	0.0	0.6	0.0	∞	0.6%
US Personal Income MoM SA	3%	-2153	0.0**	0.1**	45%	-2109	0.0**	0.2***	0.2	0.3***	3.3***	0.0	0.0***	0.0	0.0	15	3.4%
ISM Manufacturing PMI SA	1%	996	0.1	0.2*	5%	939	2.2	0.1	0.1	0.0	1.1	0.0	0.0	15.4*	0.0	5	3.7%
US Manufacturers New Orders To	1%	-1777	0.0	0.0	7%	-1638	0.0*	0.1*	-0.4**	0.1	0.1	0.0	0.0	0.0	0.0	7	3.0%
Federal Reserve Consumer Credi	1%	11472	668.9*	0.1	8%	10667	964	0	0	0	0**	-532**	286**	-26825	7	8	1.9%
Merchant Wholesalers Inventori	1%	-1908	0.0**	0.1*	9%	-1796	0.0	0.2**	-0.2	0.2	0.9	0.0	0.0	0.0	0.0	9	1.4%
US Industrial Production MOM S	7%	-2056	0.0*	0.2***	24%	-1940	0.0	0.3***	-0.4	0.4***	3.0***	0.0	0.0	0.0	0.0	3	3.7%
GDP US Chained 2009 Dollars Qo	1%	-1860	0.0	0.0*	9%	-1787	0.0**	0.1***	0.3	0.1**	1.7**	0.0	0.0	0.0	0.0	9	5.5%
US Capacity Utilization % of T	3%	-2063	0.0	0.2**	14%	-1928	0.0	0.3***	0.5*	0.0	1.9***	0.0	0.0	0.0	0.0*	5	3.2%
US Personal Consumption Expend	9%	-2414	0.0	0.1***	15%	-2248	0.0**	0.2***	0.2	0.2***	0.0	0.0	0.0*	0.0	0.0	2	3.5%
US Durable Goods New Orders In	5%	-1080	0.0	0.1***	25%	-1087	0.0	0.2***	0.8***	0.4***	2.5***	0.0	0.0***	0.1	0.0	5	3.4%
US Auto Sales Domestic Vehicle	9%	6253	114.6***	0.3***	23%	6203	-156	0***	0***	0	0***	-31	8	-5986*	1	3	1.2%
Adjusted Retail & Food Service	11%	-1426	0.0	0.2***	36%	-1475	0.0	0.2***	1.1***	0.1	4.2***	0.0	0.0	0.0	0.0***	3	3.1%
Adjusted Retail Sales Less Aut	11%	-1533	0.0	0.2***	18%	-1535	0.0	0.3***	0.5	0.2	2.0**	0.0	0.0	0.0	0.0	2	2.8%
US Durable Goods New Orders To	0%	-1055	0.0**	0.0	15%	-1072	0.0	0.1*	-0.1	0.2	1.4*	0.0	0.0***	0.0	0.0	∞	1.4%
GDP US Personal Consumption Ch	4%	-1404	0.0	0.1**	10%	-1400	0.0**	0.1*	-0.2	0.0	0.2	0.0	0.0	0.0	0.0**	3	0.5%
ISM Non-Manufacturing NMI	4%	457	0.1	0.4**	25%	444	-5.5**	0.4**	1.9*	0.1***	1.1	-0.1	-0.1**	34.8**	-0.1*	6	1.6%
US Manufacturing & Trade Inven	2%	-2247	0.0	0.1**	14%	-2152	0.0**	0.2***	-0.6**	0.0	1.5***	0.0	0.0	0.0	0.0	7	2.5%
Philadelphia Fed Business Outl	2%	1738	-0.7	0.3**	12%	1624	7.3***	0.2	-1.1*	-0.1	2.9***	-0.1	0.0	-18.2	-0.2**	6	2.5%
MNI Chicago Business Barometer	3%	1368	0.7**	0.3**	6%	1293	2.6	0.4**	0.7	0.0	2.5*	0.1	0.2	-4.2	0.0	2	2.5%
Conference Board US Leading In	6%	-2358	0.0*	0.1***	35%	-2311	0.0**	0.2***	0.2	0.2***	2.0***	0.0*	0.0**	0.0	0.0	6	2.6%
Conference Board Consumer Conf	14%	1414	0.3	0.7***	20%	1325	0.3	0.6***	0.1	0.0	2.6**	0.4*	-0.1	-0.6	-0.1	1	3.3%
US Empire State Manufacturing	1%	1272	-0.5	0.2	11%	1268	3.0	0.1	1.0	-0.1	4.3***	0.2	-0.1	-37.8	-0.3**	11	1.7%
Richmond Federal Reserve Manuf	1%	970	-1.0	0.2	18%	959	-3.2	0.2	1.5**	-0.1	3.3***	0.4	-0.2	-62.7	0.0	18	0.2%
ISM Milwaukee Purchasers Manuf	0%	451	0.0	0.0	12%	456	10.1**	0.1	-0.4	-0.2**	0.9	0.5	0.5	1.9	0.0	∞	0.1%
University of Michigan Consume	3%	2109	-0.2*	0.5***	9%	2095	1.1	0.5***	-0.3	0.0	2.1***	0.2**	-0.1**	9.1	0.0	3	2.6%
Dallas Fed Manufacturing Outlo	8%	652	-3.4***	0.6**	14%	659	0.8	0.6**	-0.7	0.0	0.3	0.1	-0.3	-52.5	-0.1	2	0.2%
US PPI Finished Goods Less Foo	9%	-1885	0.0	0.3***	18%	-1734	0.0**	0.3***	0.5	0.8***	1.1*	0.0	0.0	0.0	0.0	2	3.4%
US CPI Urban Consumers MoM SA	1%	-2559	0.0	0.0	21%	-2408	0.0**	0.2***	0.2	0.2***	1.1***	0.0	0.0	0.0	0.0	21	3.8%
US CPI Urban Consumers Less Fo	2%	-2714	0.0	-0.1**	9%	-2523	0.0**	-0.1	-0.3	-0.4**	0.4	0.0	0.0*	0.0	0.0	5	3.7%
Bureau of Labor Statistics Emp	1%	-778	0.0	0.0	10%	-718	0.0	0.0	-0.1	0.1	0.0	0.0	0.0*	0.0	0.0	10	2.8%
US Output Per Hour Nonfarm Bus	3%	-1110	0.0***	0.1**	13%	-1090	0.0	0.1***	0.1	0.1***	0.4	0.0	0.0	0.1	0.0	4	3.0%
US PPI Finished Goods SA MoM%	13%	-1554	0.0	0.2***	29%	-1533	0.0***	0.4***	1.0**	0.4***	2.0**	0.0	0.0	0.0	0.0	2	3.6%
US Import Price Index by End U	1%	-1661	0.0	0.0	23%	-1677	0.0	0.1***	0.0	0.3***	2.5***	0.0**	0.0	0.0	0.0*	23	2.0%
US GDP Price Index QoQ SAAR	9%	-1189	0.0	0.2***	41%	-1238	0.0	0.3***	-0.4**	0.0	3.5***	0.0*	0.0*	0.0	0.0	5	1.1%
US Personal Consumption Expend	1%	-1660	0.0**	-0.1	27%	-1689	0.0	0.1	-0.6**	-0.1	1.4***	0.0	0.0	0.0	0.0	27	1.3%
US Personal Consumption Expend	1%	-1524	0.0	0.0	12%	-1528	0.0	0.1*	0.1	0.0	0.6**	0.0	0.0	0.0	0.0	12	0.5%
Average	3.7%	923	-	-	17%	896	-	-	-	-	-	-	-	-	-	-	2.3%
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%	28%	70%	5%	5%	5%	5%	-	-
Popularity-weighted % of P&SC	-	-	6%	64%	-	-	8%	70%	17%	39%	75%	6%	3%	5%	2%	-	-
P&SC of most popular indicators	-	-	0%	67%	-	-	11%	67%	22%	33%	78%	11%	0%	11%	0%	-	-

Table 4: Cumulative average returns (CAR) around macroeconomic announcements.

Panel A reports the cumulative average returns (CAR) around macroeconomic announcements for positive market responses across several markets and time-frames (in minutes). Absolute CAR are reported on the left sub-panel, whereas the CAR for each time-frame as percentage of the CAR for the $[t-60min, t+1min]$ interval is reported on the right sub-panel. Panel B reports the similar CAR information but for negative market responses. t is the time of announcement of macroeconomic data releases.

Panel A - Cumulative average return (CAR) for positive market responses										
	Absolute (in Bps)					As percentage of [t-60min, t+1min]				
	[t-60, t-30]	[t-30, t-1]	[t-1, t+1]	[t-60, t+1]	[t+1,t+60]	[t-60, t-30]	[t-30, t-1]	[t-1, t+1]	[t-60, t+1]	[t+1,t+60]
S&P500	0.2	-0.9	9.9	9.2	0.1	2%	-10%	108%	100%	1%
Euro-Stoxx	-0.4	-0.5	16.3	15.5	1.6	-2%	-3%	106%	100%	10%
FTSE100	-0.2	-0.4	10.1	9.5	1.4	-2%	-5%	106%	100%	15%
2y Bund	-0.0	-0.0	1.2	1.1	0.2	-3%	-1%	104%	100%	20%
2y T-Note	-0.1	0.2	2.3	2.4	0.6	-3%	7%	96%	100%	23%
10y Gilt	-0.5	0.3	6.6	6.3	1.4	-8%	5%	103%	100%	21%
WTI Oil	1.2	0.6	6.7	8.5	-1.9	14%	7%	78%	100%	-22%
Gold	-0.3	0.8	7.1	7.7	0.3	-3%	11%	92%	100%	4%
Copper	-1.3	0.5	8.8	7.9	2.6	-17%	6%	110%	100%	33%
USDGBP	-0.5	0.0	4.1	3.6	0.4	-14%	0%	114%	100%	10%
USDEUR	-0.2	-0.1	5.4	5.1	-1.1	-3%	-1%	105%	100%	-21%
USDJPY	-0.2	0.4	5.8	6.0	0.7	-3%	7%	96%	100%	12%
USDCHF	-0.4	-0.2	5.4	4.8	0.4	-8%	-5%	113%	100%	7%
USDAUD	-0.3	-0.2	8.4	8.0	1.5	-3%	-2%	105%	100%	19%
USDCAD	0.5	0.3	5.4	6.2	-0.6	7%	5%	87%	100%	-9%
Panel B - Cumulative average return (CAR) for negative market responses										
	Absolute (in Bps)					As percentage of [t-60min, t+1min]				
	[t-60, t-30]	[t-30, t-1]	[t-1, t+1]	[t-60, t+1]	[t+1,t+60]	[t-60, t-30]	[t-30, t-1]	[t-1, t+1]	[t-60, t+1]	[t+1,t+60]
S&P500	0.1	0.1	-9.7	-9.5	-0.5	-1%	-1%	102%	100%	5%
Euro-Stoxx	1.0	0.9	-16.2	-14.2	-3.5	-7%	-7%	114%	100%	24%
FTSE100	1.2	0.6	-10.3	-8.5	-0.9	-14%	-7%	121%	100%	10%
2y Bund	0.0	-0.1	-1.2	-1.2	-0.1	0%	7%	94%	100%	8%
2y T-Note	0.0	-0.3	-2.4	-2.7	-0.3	-1%	10%	91%	100%	13%
10y Gilt	0.4	0.1	-6.8	-6.2	-0.6	-7%	-2%	109%	100%	9%
WTI Oil	-0.9	0.2	-6.8	-7.6	-0.5	12%	-2%	90%	100%	7%
Gold	-0.2	0.8	-6.9	-6.3	0.4	3%	-13%	110%	100%	-6%
Copper	0.2	-1.0	-8.8	-9.6	1.3	-2%	11%	92%	100%	-14%
USDGBP	-0.0	-0.4	-4.2	-4.7	0.3	1%	8%	91%	100%	-7%
USDEUR	0.3	-0.2	-5.3	-5.1	0.3	-5%	3%	102%	100%	-5%
USDJPY	0.3	0.5	-5.6	-4.7	-0.1	-7%	-11%	118%	100%	1%
USDCHF	0.4	-0.3	-5.2	-5.2	0.4	-7%	6%	101%	100%	-7%
USDAUD	-0.0	-0.8	-7.9	-8.8	1.0	0%	10%	90%	100%	-12%
USDCAD	0.2	-0.2	-4.9	-4.9	0.4	-4%	4%	100%	100%	-9%

Table 5: Results of restricted and unrestricted market response models per market

The table below reports results of *anchor-only* (restricted), *unrestricted* and *unrestricted-extended* market response models 1) per market evaluated, 2) aggregated per geographical coverage (i.e., local or foreign), 3) aggregated across asset classes s (i.e., stocks, bonds, FX and commodities), 4) aggregated per type of macroeconomic indicator to predict economic surprises (i.e., growth or inflation) and 5) aggregated using popularity weights. We aggregate results by averaging statistics from the individual (macro indicator-specific) models. Statistics reported are the average R^2, the explanatory power; AIC, the Akaike Information Coefficient; Coeff.>0, the percentage of positive coefficients, hit-ratios and RMSE(x1000). Hit-ratios and RMSEs are reported for our train and out-of-sample or test data, whereas other statistics are calculated in-sample. The in-sample period extends through the full data set for each indicator, whereas our out-of-sample period (our test data set) comprises of the last 25 percent of observations of the data set. The first 75 percent of the data is our training data set, which is used for tuning and estimation of predictive models. For machine learning-based predictive models, only AIC (for Ridge), hit-ratios and RMSEs are reported because inference of other statistics is not straight forward or because results may get distorted when aggregated. Statistics for the individual markets are also averages as they are aggregated from models that are based on the set of individual macroeconomic indicators investigated by us.

	Anchor-only response model					Unrestricted response model					Unrestricted-extended response model					Unrestr. Ridge response model			Unrestr. RF response model	
			%E(Surprise)	Hit-ratio	RMSE (x1000)			%E(Surprise)	Hit-ratio	RMSE (x1000)			%E(Surprise)	Hit-ratio	RMSE (x1000)		Hit-ratio	RMSE (x1000)	Hit-ratio	RMSE (x1000)
Markets	R^2	AIC	coeff. > 0	Train/Test	Train/Test	R^2	AIC	coeff. > 0	Train/Test	Train/Test	R^2	AIC	coeff. > 0	Train/Test	Train/Test	AIC	Train/Test	Train/Test	Train/Test	Train/Test
S&P500	2.0%	-1298	60%	53% / 48%	2.16 / 1.16	2.4%	-1273	65%	54% / 49%	2.16 / 1.15	36%	-1241	60%	64% / 51%	1.76 / 2.53	-1306	53% / 51%	1.44 / 1.31	54% / 50%	2.20 / 1.21
Euro-Stoxx	2.3%	-1128	63%	55% / 51%	3.28 / 1.65	3.4%	-1130	67%	55% / 52%	3.26 / 1.67	38%	-1103	53%	64% / 52%	2.66 / 2.84	-1169	53% / 52%	0.95 / 1.08	53% / 54%	3.36 / 1.80
FTSE100	2.2%	-1620	60%	54% / 49%	2.02 / 0.99	2.7%	-1589	67%	54% / 51%	2.00 / 0.97	35%	-1562	42%	63% / 52%	1.63 / 1.35	-1630	54% / 52%	1.48 / 1.55	53% / 54%	2.06 / 0.95
2y Bund	3.0%	-1133	71%	53% / 46%	0.21 / 0.09	4.3%	-1134	83%	55% / 52%	0.21 / 0.10	46%	-1094	50%	66% / 49%	0.16 / 0.20	-1192	52% / 49%	1.50 / 1.16	51% / 47%	0.22 / 0.10
2y T-Note	1.1%	-1066	52%	54% / 51%	0.43 / 0.22	2.6%	-1058	79%	54% / 52%	0.42 / 0.22	46%	-1017	50%	66% / 52%	0.33 / 0.36	-1100	54% / 52%	0.21 / 0.13	50% / 49%	0.44 / 0.23
10y Gilt	1.1%	-1499	68%	55% / 50%	1.18 / 0.90	2.2%	-1488	73%	55% / 49%	1.15 / 0.89	28%	-1445	51%	65% / 52%	0.99 / 1.10	-1445	54% / 51%	2.19 / 1.30	53% / 50%	1.18 / 0.89
WTI Oil	0.9%	-1334	58%	53% / 50%	1.47 / 1.31	1.1%	-1311	67%	52% / 49%	1.47 / 1.32	27%	-1271	58%	61% / 51%	1.25 / 1.87	-1304	51% / 51%	3.30 / 1.98	50% / 50%	1.52 / 1.32
Gold	0.8%	-1654	51%	54% / 51%	1.47 / 1.92	0.7%	-1620	42%	54% / 50%	1.48 / 1.93	24%	-1576	51%	61% / 50%	1.31 / 2.07	-1611	53% / 51%	0.97 / 1.12	52% / 51%	1.51 / 1.93
Copper	1.2%	-1311	63%	54% / 47%	1.50 / 1.08	2.4%	-1298	65%	54% / 50%	1.49 / 1.09	33%	-1261	49%	63% / 52%	1.23 / 1.36	-1326	54% / 53%	2.01 / 1.08	51% / 50%	1.52 / 1.08
USDGBP	1.0%	-2457	67%	53% / 51%	0.76 / 0.82	1.1%	-2387	67%	53% / 52%	0.76 / 0.83	17%	-2329	70%	60% / 48%	0.70 / 0.89	-2367	52% / 49%	1.48 / 1.98	52% / 50%	0.78 / 0.83
USDEUR	1.0%	-2273	63%	53% / 51%	0.96 / 1.10	1.4%	-2267	63%	55% / 52%	0.96 / 1.10	18%	-2213	56%	60% / 51%	0.87 / 1.15	-2251	53% / 52%	1.16 / 0.96	54% / 53%	0.97 / 1.10
USDJPY	1.3%	-2113	33%	53% / 50%	1.08 / 1.20	1.6%	-2114	23%	54% / 51%	1.08 / 1.21	19%	-2065	30%	61% / 50%	0.99 / 1.29	-2102	53% / 50%	0.98 / 1.12	53% / 53%	1.11 / 1.22
USDCHF	1.1%	-2154	35%	53% / 50%	0.98 / 1.09	1.4%	-2155	23%	54% / 50%	0.98 / 1.09	19%	-2106	40%	61% / 51%	0.89 / 1.17	-2144	52% / 50%	0.43 / 0.24	52% / 51%	1.00 / 1.10
USDAUD	1.8%	-1837	51%	52% / 50%	1.43 / 1.26	1.3%	-1823	44%	53% / 49%	1.43 / 1.27	24%	-1784	49%	61% / 52%	1.27 / 1.39	-1820	52% / 51%	0.77 / 0.85	52% / 52%	1.47 / 1.27
USDCAD	0.9%	-1960	44%	54% / 50%	0.95 / 1.04	0.8%	-1960	60%	53% / 48%	0.95 / 1.04	19%	-1910	49%	62% / 50%	0.86 / 1.14	-1947	53% / 50%	1.09 / 1.24	53% / 52%	0.96 / 1.05
Local (avg)	1.5%	-1182	60%	54% / 50%	1.30 / 0.69	2.5%	-1165	65%	54% / 51%	1.29 / 0.68	41%	-1129	55%	65% / 52%	1.05 / 1.44	-1203	53% / 51%	1.31 / 0.77	52% / 50%	1.32 / 0.72
Foreign (avg)	1.4%	-1729	56%	54% / 50%	1.33 / 1.11	1.9%	-1714	57%	54% / 50%	1.32 / 1.12	27%	-1671	50%	62% / 51%	1.14 / 1.37	-1716	53% / 51%	1.33 / 1.20	52% / 51%	1.36 / 1.13
Stocks (avg)	2.2%	-1349	61%	54% / 50%	2.49 / 1.27	2.8%	-1331	67%	54% / 51%	2.47 / 1.26	36%	-1302	52%	64% / 52%	2.02 / 2.24	-1368	53% / 52%	2.50 / 1.45	54% / 53%	2.54 / 1.32
Bonds (avg)	1.8%	-1232	64%	54% / 49%	0.61 / 0.41	3.0%	-1227	78%	55% / 51%	0.59 / 0.40	40%	-1185	50%	66% / 51%	0.50 / 0.55	-1246	53% / 51%	0.60 / 0.44	51% / 49%	0.61 / 0.41
Currencies (avg)	1.2%	-2132	49%	53% / 51%	1.03 / 1.08	1.3%	-2118	47%	54% / 50%	1.03 / 1.09	20%	-2068	49%	61% / 50%	0.93 / 1.17	-2105	52% / 50%	1.03 / 1.12	52% / 52%	1.05 / 1.09
Commodities (avg)	1.0%	-1433	57%	54% / 49%	1.48 / 1437.32	1.4%	-1410	58%	53% / 50%	1.48 / 1.45	28%	-1370	53%	62% / 51%	1.27 / 1.77	-1414	52% / 52%	1.49 / 1.56	51% / 50%	1.52 / 1.45
Growth (avg)	1.5%	-2623	60%	54% / 50%	1.35 / 1.09	2.0%	-2608	69%	54% / 51%	1.35 / 1.10	29%	-2561	57%	63% / 51%	1.16 / 1.44	-2600	53% / 51%	1.36 / 1.17	52% / 51%	1.38 / 1.11
Inflation (avg)	1.4%	-1522	48%	53% / 50%	1.16 / 0.88	1.8%	-1504	48%	54% / 49%	1.15 / 0.88	31%	-1464	44%	63% / 50%	0.96 / 1.20	-1535	53% / 50%	1.15 / 0.98	52% / 51%	1.18 / 0.90
Average	1.5%	-1656	56%	54% / 50%	1.32 / 1.06	2.0%	-1640	59%	54% / 50%	1.32 / 1.06	29%	-1599	50%	63% / 51%	1.12 / 1.40	-1648	53% / 51%	1.33 / 1.14	52% / 51%	1.35 / 1.07
Pop-weighted avg.	1.4%	-1658	58%	54% / 50%	1.39 / 1.13	2.0%	-1639	68%	54% / 50%	1.38 / 1.13	26%	-1598	56%	62% / 51%	1.20 / 1.40	-1640	52% / 51%	1.39 / 1.21	52% / 51%	1.41 / 1.15
Most pop. indicators (avg)	1.1%	-1246	63%	54% / 50%	1.52 / 1.23	2.1%	-1691	79%	54% / 51%	1.51 / 1.23	24%	-1651	59%	62% / 52%	1.34 / 1.44	-1687	52% / 52%	1.53 / 1.31	52% / 50%	1.54 / 1.26

Table 6: Results for unrestricted-extended market response models per factor

Panel A reports details on the fit of unrestricted-extended models for the in-sample period. Panel B reports details on the fit of unrestricted-extended models for the out-of-sample period. Across the two panels, we contrast results from the *Unrestricted-extended*, *Unrestricted-Ridge regression* and *Unrestricted-Random Forest* models to provide some interpretation into our results. As the three models used do not provide a common variable for direct comparison, these statistics are mostly used here to map the best predictors and, perhaps, found confirmation of that from other models.

	Panel A - Unrestricted reponse models - Train set				Panel B - Unrestricted reponse models - Test set/Out-of-sample			
	Extended (OLS)		Ridge	Random Forest	Extended (OLS)		Ridge	Random Forest
	% positive	% signifcant	% positive	Importance (x10$\hat{5}$)	% positive	% signifcant	% positive	Importance (x10$\hat{5}$)
Intercept	49%	10%	47%	-	50%	10%	46%	-
UnES	58%	13%	65%	2.4	56%	12%	66%	2.0
ESA	52%	11%	59%	1.4	51%	10%	57%	1.2
Skew	48%	10%	62%	2.3	45%	10%	52%	1.9
Std	44%	13%	50%	1.6	45%	12%	47%	1.2
SurvLag	52%	10%	47%	1.6	51%	9%	49%	1.3
Inflation	52%	14%	52%	1.8	57%	12%	56%	1.4
Growth	51%	8%	47%	1.8	51%	7%	50%	1.4
Stocks	66%	28%	64%	3.5	65%	25%	65%	3.0
VIX	49%	13%	51%	3.4	50%	11%	53%	2.9
VIXdif	58%	15%	49%	1.9	55%	15%	55%	1.6
ret60	51%	15%	46%	1.8	51%	15%	49%	1.4
ret55	50%	19%	50%	2.1	52%	16%	54%	1.7
ret50	46%	21%	49%	2.2	47%	16%	47%	1.8
ret45	48%	18%	49%	2.3	48%	17%	49%	1.9
ret40	43%	16%	43%	2.4	45%	13%	43%	1.9
ret35	53%	16%	51%	2.7	52%	16%	52%	2.3
ret30	44%	15%	41%	2.2	42%	12%	42%	1.8
ret25	51%	17%	47%	2.2	51%	15%	51%	1.9
ret20	48%	17%	47%	2.1	50%	13%	47%	1.8
ret15	40%	21%	38%	2.5	45%	15%	41%	2.1
ret10	49%	22%	48%	2.4	53%	18%	52%	2.0
ret5	66%	27%	59%	3.5	65%	23%	65%	2.9
Average	51%	16%	51%	2.3	51%	14%	52%	1.9

Table 7: Test of market response in presence of regret

This table reports results of our test for the presence of regret within market responses, see Eqs. (11a) to (11d) in an univariate setting. $Skew_t^+$ is the skewness in forecasts when it is positive and $Skew_t^-$ when it is negative. S_t^+ is the realized surprise when it is positive and S_t^- when it is negative. Explanatory variables in these models are interaction terms between surprises and skewness in forecasts. $Skew_t^+ * S_t^-$ and $Skew_t^- * S_t^+$ are scenarios in which the expected surprise failed to forecast the realized economic surprise. $Skew_t^+ * S_t^+$ and $Skew_t^- * S_t^-$ are scenarios in which the expected surprise was successful in forecasting the economic surprise. We use Newey-West adjustments to compute coefficient standard errors. The asterisks ***, **, and * indicate significance at the one, five, and ten percent level, respectively.

	Panel A - Univariate							
Markets	R^2	Skw+Surp-	R^2	Skw+Surp+	R^2	Skw-Surp-	R^2	Skw-Surp+
S&P500	0.18%	-19.4	0.01%	5.5	0.10%	-15.9	0.07%	12.6
Euro-Stoxx	0.04%	-6.4	0.04%	8.6	0.01%	-3.0	0.12%	11.5
FTSE100	0.10%	-11.8	0.00%	3.2	0.00%	-0.7	0.00%	3.0
2y Bund	0.54%	-26.9***	0.03%	9.4	0.01%	3.0	0.07%	12.6
2y T-Note	0.35%	-3.6*	0.02%	0.8	1.50%	7.4***	0.00%	0.2
10y Gilt	0.07%	-14.5	0.09%	-25.6	0.07%	13.9	0.00%	-0.1
WTI Oil	0.14%	-29.0	0.05%	-24.4	0.06%	20.3	0.01%	9.6
Gold	0.00%	-1.9	0.01%	-3.5	0.01%	-2.5	0.08%	9.2
Copper	0.07%	-12.9	0.02%	-9.0	0.01%	5.8	0.00%	-2.8
USDGBP	0.11%	-13.8	0.04%	12.5	0.09%	15.6	0.02%	8.3
USDEUR	0.00%	-2.1	0.14%	17.9	1.37%	38.6***	0.34%	22.7**
USDJPY	0.28%	-16.4**	0.01%	3.0	0.03%	-5.8	0.21%	15.2*
USDCHF	0.05%	3.4	0.00%	0.4	0.06%	3.2	0.09%	4.6
USDAUD	0.08%	-6.7	0.04%	-7.4	0.00%	1.8	0.13%	9.8
USDCAD	0.04%	-7.1	0.00%	3.2	0.12%	-13.1	0.01%	3.0
% coefficients > 0		7%		67%		60%		87%
% sig. coefficients > 0		0%		-		100%		100%

Table 8: Aggregated results of anchor-only and unrestricted economic surprise models for Europe, the UK and Japan

Panel A reports the percentage of statistical significant coefficients (factors) across anchor-only and unrestricted regression models for economic surprises of macroeconomic 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 percent of such the ESA factor across the individual regressions run for the European macroeconomic indicators are statistically significant at the 10 percent level. Panel B reports the mean, median and standard deviation of the explanatory power (R^2) achieve by 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 macroeconomic indicators.

Region	Cont. Europe		UK		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
Skew		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
IV		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
Skew		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
IV		0.26		0.42		0.19

Table 9: Results of two-component (expected- and unexpected-economic surprise) response models

The Panel A of the table below reports aggregated results for the *two-component response model* based on the anchor-only and on the unrestricted model for economic surprises. *Two-component response models* separate economic surprises into expected surprises, as forecasted by economic surprise predictive models, and unexpected surprises, as the residual between economic surprises and expected surprises as given by Eq. (13). Statistics reported for two-component models are R^2 and the percentage of regression coefficients that are statistically significant across different markets. Panel B reports R^2 for the aggregated (univariate) *unexpected-surprise response model* as given by Eq. (14) as well as the absolute and percentage gain in R^2 delivered by *Two-component response models* versus *unexpected-surprise response models*. Regression results are reported aggregated by averaging across the 43 US economic indicators used. We use Newey-West adjustments to compute coefficient standard errors.

	Panel A - Two-component response models						Panel B - Comparison between R^2 from two-component response models and unexpected response models					
	Anchor-only two-component response model			Unrestricted two-component response model			Unexpected response models Anchor-only model			Unrestricted model		
	R^2	Expected % Significant	Unexpected % Significant	R^2	Expected % Significant	Unexpected % Significant	R^2	R^2 gain by two-component	% R^2 gain by two-component	R^2	R^2 gain by two-component	% R^2 gain by two-component
Average	9.7%	22%	62%	10.0%	32%	59%	8.8%	0.9%	49%	8.7%	1.4%	51%
Popularity-weighted average	11.6%	26%	72%	12.0%	37%	68%	10.6%	1.0%	33%	10.5%	1.5%	42%
Most pop. indicators (avg)	19.0%	26%	95%	20.1%	48%	94%	18.1%	0.9%	5%	18.4%	1.7%	9%

A) Positive responses

B) Negative responses

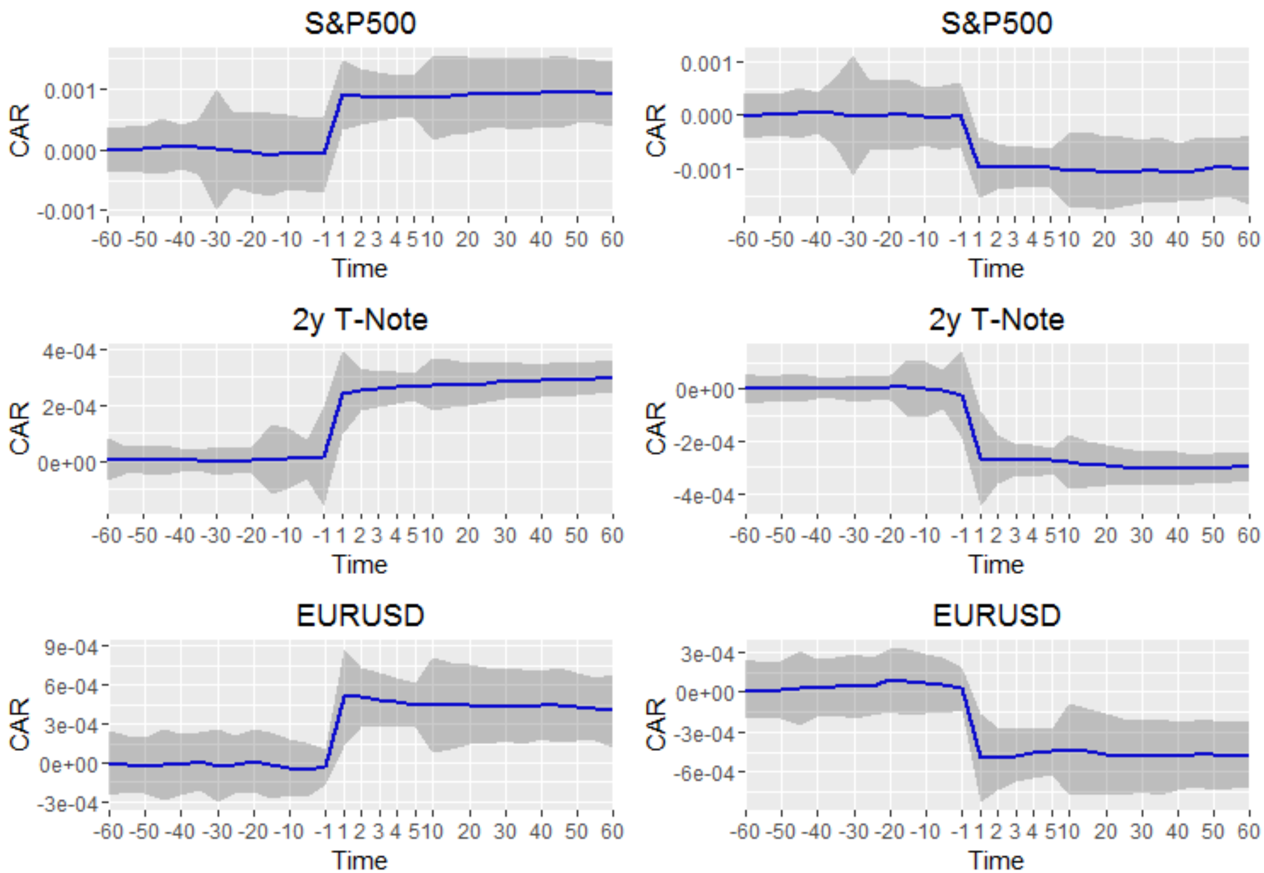
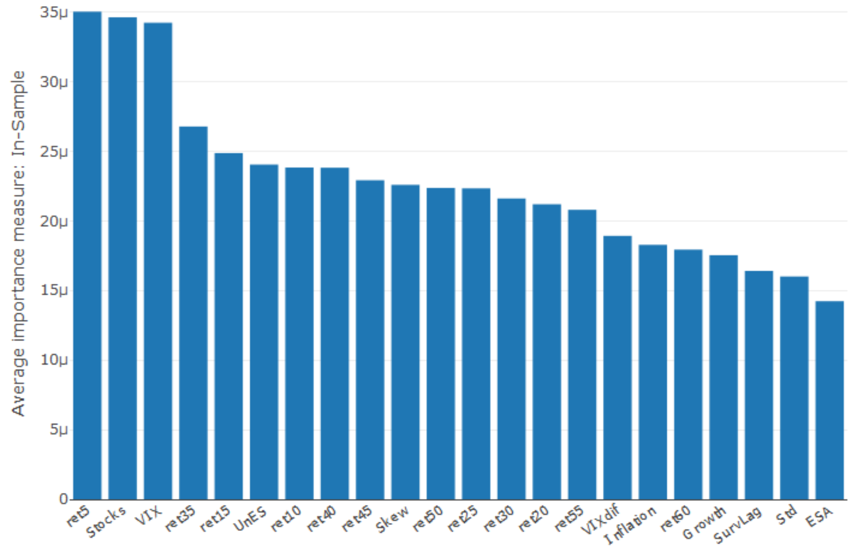


Figure 1: Cumulative average returns (CAR) around the macroeconomic announcements. The line plots above depict the CAR (across all indicators) around the time of macroeconomic announcements (in blue). The time of macro economic announcements within these plots is $t = 0$ within the x-axis. The -60, -50, -40, -30, -20, -10, -5, -1 time-frames, which proceeds $t = 0$, represent the minutes prior to the macro announcement. The 1, 2, 3, 4, 5, 10, 20, 30, 40, 50, 60 time-frames represent the minutes after the macro announcement. The shadowed area around the CAR line show its one standard deviation (68.27^{th} percent) confidence interval. The securities evaluated are the S&P500 index future, the 2-year treasury bond future and EURUSD forwards, respectively, reported in rows one, two and three.

A) In-sample



B) Out-of-sample

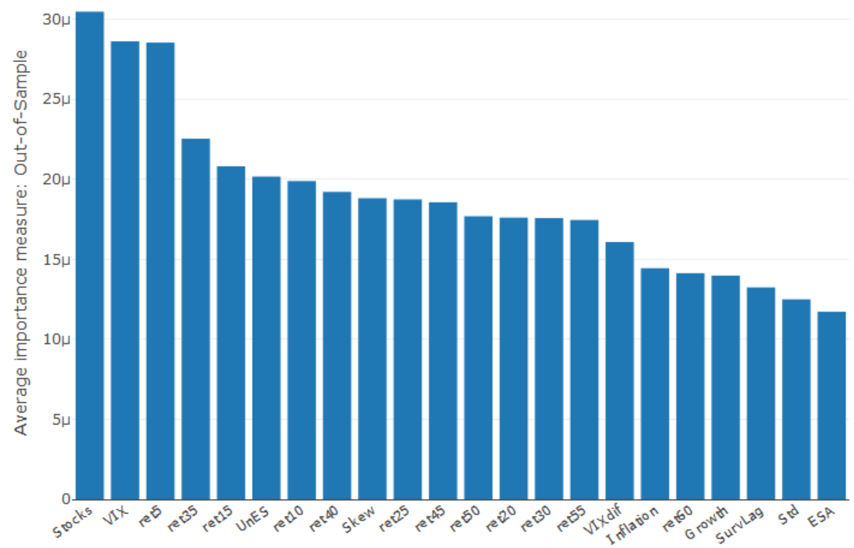


Figure 2: **Market responses .** The bar plots above depict the *Importance* measure produced by the Random Forest model applied to predict market responses around the announcements of macroeconomic data, in the train and test data sets. More specifically, the *Importance* measure computes the average node impurity across all trees grown by the Random forest, which reflects how partitions made by the different explanatory variables at each node into two sub-regions perform versus a constant fit over the entire region, i.e., a 'pure' node. For the case of regression models, performance is calculated in terms of squared errors (see Appendix A for additional details on the *Importance* measure).

A) All macro releases

B) Inflation-only releases

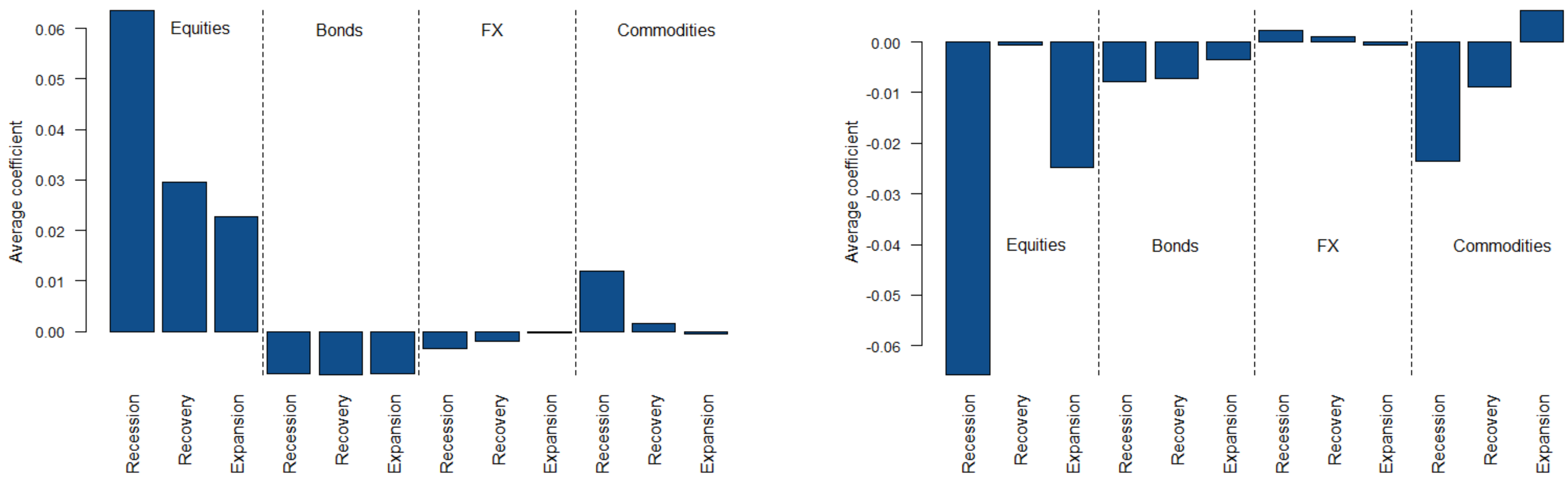
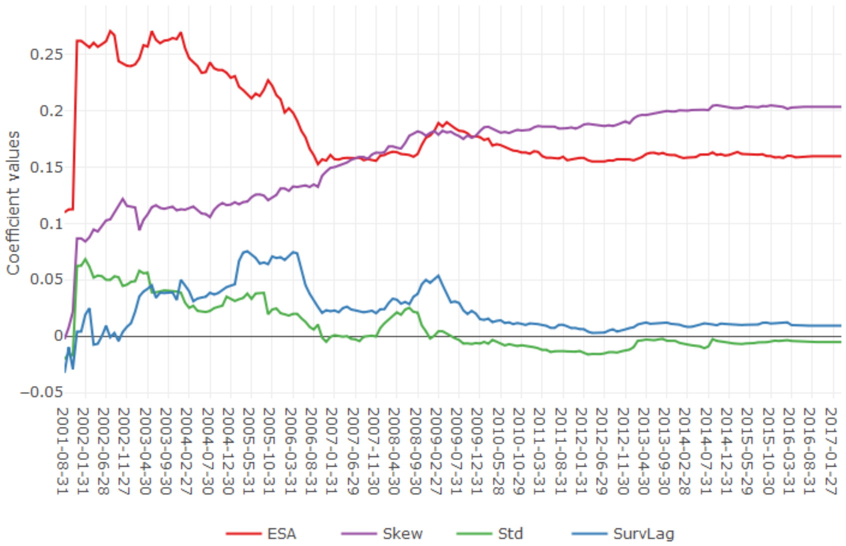


Figure 3: Market responses through the economic cycle. The bar plots above depict the average coefficient values of a univariate regression that explains responses of equities, bonds, FX and commodities markets using economic surprise (expressed in standard deviations versus the median) as explanatory variable across the three stylized phases of the economic cycle: recession, recovery and expansion. Panel A reports average coefficient values for these regression when all type of macroeconomic releases are considered. Panel B reports average coefficient values for these regression when only inflation releases are considered. Recessionary periods are identified following the classification of the US National Bureau of Economic Research (NBER). Recoveries are non-recessionary periods for which the US core inflation (CPI) is below two percent, whereas expansion is categorized as non-recessionary period for which core inflation is above two percent.

A) Coefficient values



B) P-values

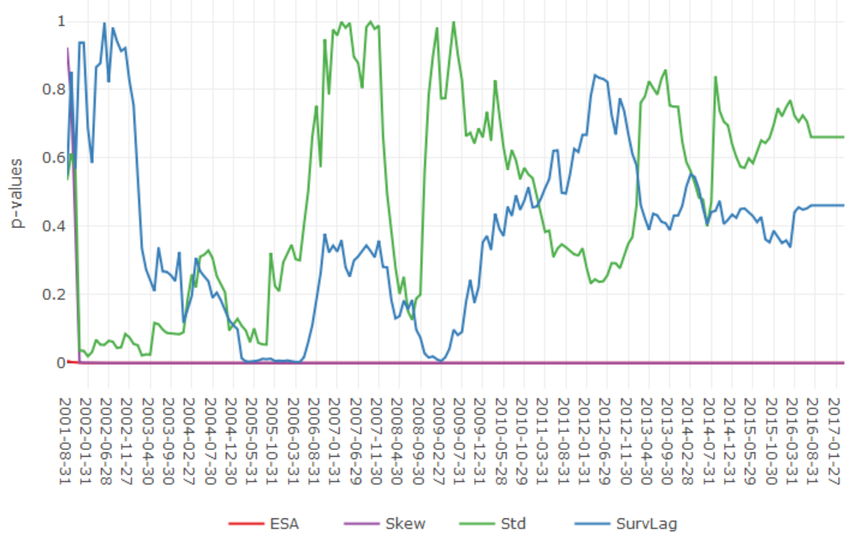


Figure 4: Economic surprise models through time. The line plots above depict the coefficient values and p-value of predictors ESA, $Skew$, Std and $SurvLag$ through time, respectively in boxes A and B. The coefficient of ESA and $Skew$ are positive and statistically significant, with few exceptions, whereas the coefficients for Std and $SurvLag$ fluctuate between a positive or a negative sign, beyond being mostly not statistically significant.

Recent Issues

All CFS Working Papers are available at *www.ifk-cfs.de*.

No.	Authors	Title
595	Renée Adams, Roman Kräussl, Marco Navone, and Patrick Verwijmeren	*Is gender in the eye of the beholder? Identifying cultural attitudes with art auction prices*
594	Amy Whitaker and Roman Kräussl	*Blockchain, Fractional Ownership, and the Future of Creative Work*
593	Petrus Ferreira, Roman Kräussl, Wayne R. Landsman, Maria Nykyforovych, and Peter Pope	*Reliability and Relevance of Fair Values: Private Equity Investments and Investee Fundamentals*
592	Volker Brühl	*Clearing of euro OTC derivatives – a milestone on the road to a Capital Markets Union in Europe*
591	Günther Gebhardt and Zoltán Novotny-Farkas	*Comparability and predictive ability of loan loss allowances – The role of accounting regulation versus bank supervision*
590	Enikő Gábor-Tóth and Dimitris Georgarakos	*Economic Policy Uncertainty and Stock Market Participation*
589	Christos Koulovatianos, Jian Li, and Fabienne Weber	*Market Fragility and the Paradox of the Recent Stock-Bond Dissonance*
588	Volker Brühl	*Clearing of euro OTC derivatives post Brexit - an analysis of the present cost estimates*
587	Salomon Faure and Hans Gersbach	*Loanable Funds vs Money Creation in Banking: A Benchmark Result*
586	Roman Goncharenko, Steven Ongena, and Asad Rauf	*The Agency of CoCo: Why Do Banks Issue Contingent Convertible Bonds?*