COGNITIVE BIASES IN INDUSTRY FORECASTS OF USDA REPORTS: IMPLICATIONS FOR PRICE REACTIONS AND MARKET SURPRISE MEASURES

Berna Karali

University of Georgia, Athens, GA, USA

Olga Isengildina-Massa

Virginia Polytechnic Institute and State University, Blacksburg, VA, USA

Scott H. Irwin

University of Illinois at Urbana-Champaign, Urbana, IL, USA

Corresponding author: Berna Karali; E-mail: bkarali@uga.edu

Cognitive Biases in Industry Forecasts of USDA Reports: Implications for Price Reactions and Market Surprise Measures

Abstract

We investigate the impact of cognitive biases and heterogeneity in firm-level forecasts of USDA corn and soybean production estimates on market price reactions using two approaches. First, after adjusting for cognitive biases, we decompose market surprises—defined as the difference between USDA estimates and private forecasts—into expected and unexpected components to test whether futures prices and volatility respond only to the unexpected component, a condition indicative of market efficiency. Second, we construct a range of market surprise measures to determine which best explains price movements on USDA report release days. We find that the corn futures market exhibits informational inefficiency with respect to anchoring bias, as prices respond to both components of the surprise, and the soybean futures market demonstrates informational efficiency, with prices reacting solely to the unexpected surprise. The pattern is reversed for attribution bias, with corn volatility reacting only to the unexpected surprise, and soybean volatility to both components of the surprise. For modeling, we find that accounting for heterogeneity among individual forecasts enhances the explanatory power of price reaction models.

Key Words

anchoring, attribution, behavioral bias, crop production, informational efficiency

JEL Classification

C33, D80, D84, E37, Q11

Cognitive Biases in Industry Forecasts of USDA Reports: Implications for Price Reactions and Market Surprise Measures

Introduction

The value, impact, and welfare effects of the U.S. Department of Agriculture (USDA) reports are frequently measured by the commodity price and volatility movements to their release (e.g., Milonas 1987; Sumner and Mueller 1989; Fortenbery and Sumner 1993; Adjemian 2012; Karali 2012; Dorfman and Karali 2015; Ying, Chen, and Dorfman 2019; Adjemian and Irwin 2020; McKenzie and Ke 2022).[1] These price movements are especially pronounced when the USDA figures deviate from market expectations, where the differences between the two are commonly referred to as market surprises (e.g., Colling, Irwin, and Zulauf 1996; Garcia et al. 1997; Good and Irwin 2006; McKenzie 2008; Karali et al. 2019). Historically, studies have relied on industry polls aggregated by news agencies like Bloomberg or Reuters—using the mean or the median of individual forecasts as proxies for market expectations. However, this approach has some limitations. First, it ignores the heterogeneity among forecasters, which could increase the richness of the analysis and the scope of the research. Fernandez-Perez et al. (2019), for instance, show that the dispersion among analysts' forecasts significantly affects bid-ask spreads in corn, soybean, and wheat futures prices during USDA report releases.

Second, it ignores various forms of biases found in market surprises. For example, Karali, Irwin, and Isengildina-Massa (2020) demonstrate that market surprises constructed using the median of firm-level expectations are subject to attenuation bias, and thus, the true price

[1] For a comprehensive review of previous studies on the impact of various USDA reports, see Isengildina-Massa, Karali, and Irwin (2024).

reactions are underestimated in previous studies.[2] Another recent study documents the existence of cognitive biases, such as anchoring and attribution, in industry expectations of corn and soybean production estimates for the USDA's August Crop Production report (Karali, Isengildina-Massa, and Irwin 2025).[3] The authors argue that market surprises become partly predictable when there is any form of systematic bias in industry expectations. The predictability of the market surprise, then, would provide profit opportunities to financial market participants who are aware of these biases, thereby altering welfare impacts.

In this study, we investigate the informational efficiency of futures markets by testing whether prices and volatility react to the predicted component of the USDA news induced by anchoring and attribution biases. To this end, we first summarize the evidence of anchoring and attribution biases in the industry expectations demonstrated in Karali, Isengildina-Massa, and Irwin (2025). We then follow the methods outlined in Campbell and Sharpe (2009) and Isengildina-Massa, Karali, and Irwin (2017) to decompose the market surprise into anticipated and unanticipated components after accounting for these biases and model the price and volatility reactions to USDA reports as a function of the decomposed surprise measures. If market participants are aware of and account for the biases, prices and volatility should only respond to the unexpected component of the surprise, implying an informationally efficient market. On the other hand, if market participants are unaware of or fail to account for the cognitive biases in industry forecasts, prices and volatility would also react to the expected component of the surprise, suggesting an informationally inefficient market.

---

[2] Attenuation bias refers to the systematic underestimation of the true relationship between variables, typically caused by measurement error in the explanatory variable.

[3] Anchoring bias describes the tendency to rely heavily on an initial, easily accessible reference point when forming forecasts, with subsequent adjustments made relative to that anchor. Attribution bias reflects analysts' overconfidence in their forecasting abilities, often stemming from previous successful predictions.

These inefficiency concerns also raise the question of which surprise measure should be used to assess market reaction to USDA reports. We tackle this question by developing various surprise measures that incorporate the heterogeneity among private forecasts, as well as the predictability due to cognitive biases, and comparing the explanatory power of the price reaction tests across different specifications.

We find that market participants fail to account for anchoring bias in corn forecasts, as prices respond to the predicted component of surprises—even in ex-ante analyses using the latest available information. However, corn price volatility reacts only to the unpredicted component, suggesting awareness of attribution bias. In contrast, soybean prices respond only to unpredicted surprises, indicating that anchoring bias is accounted for, while volatility responds to the predicted component, implying attribution bias is overlooked. In terms of modeling, the highest explanatory power for corn comes from using the median forecast to compute surprises and including the interquartile range of firm-level forecasts in an ordinary least squares (OLS) model. For soybeans, the best results come from decomposed surprises in a fixed-effects panel model, incorporating the median absolute deviation of firm-level forecasts. Overall, our results highlight that accounting for firm-level forecast heterogeneity improves the explanatory power of price reactions to USDA news. The findings of this study will help us improve how we incorporate industry expectations information into market reaction tests to assess the value and welfare effects of USDA reports.

Empirical Framework

Based on the findings in Karali, Isengildina-Massa, and Irwin (2025), we focus on anchoring and attribution biases in industry forecasts to investigate whether financial market participants are informationally efficient by testing price and volatility reactions to the predicted surprise components induced by these cognitive biases. Note that although the regression equations testing for anchoring and attribution biases use the forecast error and absolute forecast error, respectively, as dependent variables, the forecast error itself is calculated in the same way as the market surprise commonly employed in the literature. For brevity, we focus on the anchoring bias in the main body of the paper and present the methods and empirical results for the attribution bias in the appendix.

Anchoring bias and price reaction tests

Anchoring bias is defined as a form of cognitive bias in which people form their forecasts by starting from an easily available reference point and then make adjustments based on this value (Tversky and Kahneman 1974). Karali, Isengildina-Massa, and Irwin (2025) follow Campbell and Sharpe (2009) and test for anchoring bias using a regression equation with the forecast error as the dependent variable. We estimate the following regression equation with firm fixed effects:

$$(1) \quad FE_{it} = \alpha_i + \varphi \; Deviation\ from\ Anchor_t + \varepsilon_{it} \; ,$$

where $i = 1, 2, \cdots, N$ represents the firms making forecasts and $t = 1, 2, \cdots, T$ denotes the years. The dependent variable FE_{it} is the forecast error (i.e., surprise), calculated in percentages to account for the changes in magnitudes across years as follows:

$$(2) \quad FE_{it} = 100 \times \frac{(Actual_t - Forecast_{it})}{Actual_t} \; ,$$

where $Actual_t$ is the actual value in year *t* and $Forecast_{it}$ is the forecast made by firm *i* in year *t*. The deviation from anchor is defined in percentages as $100 \times [(Forecast_{it} - Anchor_t)/Anchor_t]$ and $Anchor_t$ is the initial starting point. Forecast errors are systematically biased in a predictable manner consistent with anchoring if the slope estimate in equation (1) is positive and statistically significant.

The finding of an anchoring bias makes the forecast error, which is equivalent to the market surprise measure used in the literature, partly predictable. The predicted dependent variable from equation (1), $\widehat{FE}_{it}$, can be considered as the expected surprise, and the predicted residuals, $\hat{\varepsilon}_{it} = FE_{it} - \widehat{FE}_{it}$, as the unexpected, "true" surprise (Campbell and Sharpe 2009; Isengildina-Massa, Karali, and Irwin 2017). We can then use the decomposed surprise measures to test for the price reactions by estimating the following regression equation with firm fixed effects:

$$(3) \quad \Delta P_{it} = \omega_i + \eta_1 \widehat{FE}_{it} + \eta_2 \, \hat{\varepsilon}_{it} + e_{it},$$

where $\Delta P_{it} = 100 \times (\ln P_t - \ln P_{t-1})$ is the continuously compounded daily return on the commodity futures contract with price P_t, and it is the same for all firms ($\Delta P_{it} = \Delta P_t, \forall i$). The null hypothesis of an informationally efficient market is $\eta_1 = 0$, which suggests that market participants are aware of the anchoring bias in industry forecasts and that this information is already reflected in prices. To account for the increased sampling variability induced by using generated regressors, we estimate the model in (3) via the generalized method of moments (GMM) outlined in Campbell and Sharpe (2009).

Data

We utilize the same data set in Karali, Isengildina-Massa, and Irwin (2025), containing firm-level forecasts for corn and soybean production in USDA's Crop Production reports. These reports are prepared and published by the National Agricultural Statistical Service (NASS) agency of USDA and contain survey-based estimates of yield and production for major crops consistent with their growing cycles. USDA publishes the first marketing-year production estimates for corn and soybeans in August, then revises them through November, and finalizes them in the January Crop Production Annual Summary (CPAS) report. We attain futures price data from CRB Trader. Corn and soybean futures contracts are traded at the Chicago Mercantile Exchange (CME) Group, with multiple contract maturities. We use the December contract for corn and the November contract for soybeans to represent the new crop futures series.

The proprietary firm-level forecast dataset includes companies that participate in polls conducted by news wire agencies and provide forecasts for crop production, crop yield, planted acreage, and stocks for corn, soybean, and wheat varieties.[4] Because of the importance of being the first production estimates for corn and soybeans, as well as data availability, we focus on the August Crop Production report. Following Karali, Isengildina-Massa, and Irwin (2025), we take the production figures in the August report as the actual values firms try to forecast.[5]

Our sample of industry forecasts for the upcoming August Crop Production report covers 1992 to 2021.[6] The dataset is an unbalanced panel, with some companies disappearing from the sample in the early 2000s (the earliest in 2002) and some entering the sample late (the latest in 2012). We exclude the firms without forecasts after 2010 to avoid using a panel dataset in which

---

[4] Wire news services, such as Bloomberg and Reuters, poll advising companies, commodity market experts, and market analysts for their expectations regarding the upcoming USDA reports. These are also known as trade estimates. While the poll results are available to Bloomberg subscribers, to the best of our knowledge, they only go back to 2010. In contrast, our private source has been compiling industry polls from various sources since the late 1980s, providing us with a unique opportunity to leverage industry forecasts that extend further back into the past.

[5] As discussed in Karali, Isengildina-Massa, and Irwin (2025), Bloomberg has started releasing two sets of forecasts in recent years. One set includes analysts' expectations of USDA figures in the upcoming reports, and the other contains analysts' forecasts for the unobservable actual value. However, these two sets of forecasts are not available for each report, and for a given report, such as Grain Stocks, their availability varies by the report month. Since our proprietary data set includes only one forecast for each firm without indicating to which forecast it refers, we assume that the firms in our dataset forecast the crop production figures in August reports.

[6] Even though the industry forecast dataset begins in 1989, the number of firms providing forecasts is sparse before 1992, constraining the start of our sample period. Some companies disappear from the dataset in the early 2000s (the earliest in 2002), while others enter the dataset relatively late (the latest in 2012).

some firms do not overlap with others or have only one common time period. From the remaining sample, we exclude firms with fewer than ten observations between 1992 and 2021, resulting in an unbalanced panel of 24 firms. The number of firms per year ranges from nine to 24, with an average of 15.3 firms over the 30-year period.

Figure 1 displays the empirical probability distributions (i.e., kernel) of forecast errors (i.e., surprises) and absolute forecast errors (i.e., absolute surprises) relative to a normal distribution. Soybean forecast errors are distributed relatively widely, indicating larger standard deviations. Figure 2 presents the distributions of forecast errors and absolute forecast errors for each year across firms that provided a forecast. The horizontal lines in each box represent the median, and the dots represent the "outliers." For corn, the median absolute forecast error is lowest in 2001 and highest in 2019, and for soybeans, it is lowest in 2001 and highest in 2015.

Empirical Results

Anchoring bias

We assign four different measures for the anchor when estimating equation (1), following Karali, Isengildina-Massa, and Irwin (2025). The first anchoring variable is the final production of the previous crop year published in the USDA's January CPAS report, $Anchor\ 1_t = Final\ Production_{t-1}$. To account for the adjustment made by the forecasters to smooth out the fluctuations in production across years, we take the second anchoring variable as the average of the final production values in the last three years, $Anchor\ 2_t = \frac{1}{3}\sum_{j=1}^{3} Final\ Production_{t-j}$. While the previous year's final production can be considered as the latest information the forecasters have about the upcoming crop year's production, some firms might think there is a pattern in USDA's estimates for the first production figures for the crop year and therefore use the actual value in the previous year's August Crop Production report as a starting point. Accordingly, we take the third anchoring variable as $Anchor\ 3_t = Actual_{t-1}$, and the fourth anchoring variable is the average of the actual values in the last three periods to smooth out the fluctuations over time, $Anchor\ 4_t = \frac{1}{3}\sum_{j=1}^{3} Actual_{t-j}$.

The regression results are presented in table 1 for corn and in table 2 for soybeans.[7] As indicated in the previous study of Karali, Isengildina-Massa, and Irwin (2025), corn forecasts exhibit reverse anchoring, evidenced by the negative coefficient estimates across all specifications. Thus, firms make adjustments in the opposite direction of their initial starting value (i.e., anchor). The model with the lowest Akaike information criteria (AIC) and Bayesian information criteria (BIC) is indicated with the darker shade, and the model with the second lowest AIC and BIC is indicated with the lighter shade. It appears that the models with an anchoring variable created with smoothing fit the data better.

For soybeans, coefficient estimates are positive across all specifications, but are statistically significant only in models with smoothed anchoring variables. Thus, soybean forecasts exhibit anchoring bias, where firms base their current forecasts on a reference value (smoothed production values over the last three years) and make sufficient adjustments. Similar to corn, the model with the anchoring variable set to the average August production over the last three years has the smallest AIC and BIC, while the model with the average of the previous three years' final production as the anchoring variable has the second smallest AIC and BIC.

[7] These results replicate tables 12 and 13 in Karali, Isengildina-Massa, and Irwin (2025), with slight differences in the corn results arising from our correction of two data points misrepresenting the actual USDA figure in 1997 and 2016.

Price reaction tests with anchoring bias

The statistical evidence of anchoring and reverse anchoring bias suggests that forecast errors (i.e., market surprises) are partly predictable. If markets are informationally efficient, prices should not react to the predicted component of the surprise. The price reaction test results from equation (3) are presented in table 3 for corn and in table 4 for soybeans. For comparison purposes, column (1) of these tables presents the results obtained from the ordinary least squares (OLS) regression of futures price changes on the surprise variable taken as the median of firm-level surprises within a year. Column (2) of both tables uses firm-level surprises in a panel regression with firm fixed effects. Columns (3) and (3') decompose the surprise measure into expected and unexpected components by using the predicted residuals from the anchoring regression with the lowest AIC and BIC (columns (4) in table 1 and 2 for corn and soybeans, respectively). Similarly, columns (4) and (4') decompose the surprise measure by using the predicted residuals from the anchoring regression with the second-lowest AIC and BIC (columns (2) in tables 1 and 2).

The difference between (3) and (3'), and between (4) and (4'), stems from how the anchoring regression is estimated. In columns (3) and (4), the anchoring regression equation is estimated using the entire sample, thereby making the decomposition of the surprise ex-post. To account for the latest information about the anchoring bias in industry forecasts, we estimate the anchoring regression using a 10-year rolling window. Specifically, starting from 2002, we use the previous 10 years of data (1992-2001) to estimate equation (1) and calculate the predicted forecast errors and predicted residuals, which represent, respectively, the expected and unexpected components of the surprise in 2002. We repeat this process for each year from 2002 to 2021 to create a series of decomposed surprise measures. Columns (3') and (4') present the results obtained with ex-ante decomposed surprise measures. Recall that we estimate models (3) and (4') using the GMM approach outlined in Campbell and Sharpe (2009), which accounts for the increased sampling variability resulting from the generated regressors. For these models, we also test whether prices react to expected and unexpected surprises symmetrically (Campbell and Sharpe 2009).

For corn, the surprise coefficient estimates are negative across all models in table 3, suggesting that positive supply surprises (i.e., higher-than-expected production) lower futures prices, and negative supply surprises (i.e., lower-than-expected production) put upward pressure on prices. We find that prices react to both predicted and unpredicted surprise components, even after decomposing the surprise measure ex-ante. These results suggest that financial market participants are either unaware of or fail to account for the anchoring bias in corn production forecasts, implying an informationally inefficient market. The symmetric price reaction to both predicted and unexpected surprises is rejected, except for the model with ex-ante decomposed surprises (column (3')). When the symmetry is rejected, the magnitude of price reaction to the true surprise is about half of that to the predicted surprise.

The surprise coefficient estimates for soybeans in table 4 are negative when statistically significant, aligning with our expectations. We find that once we account for the latest information set that market participants could use to adjust for the anchoring bias (i.e., using ex-ante decomposed surprises), prices only react to the true surprise. The coefficient estimate for the predicted surprise is statistically insignificant in both (3') and (4'). These findings reveal that market participants are aware of and account for the anchoring bias in industry forecasts of soybean production, implying an informationally efficient market.

Volatility reaction tests with attribution bias

We estimate the models in Karali, Isengildina-Massa, and Irwin (2025) to determine the factors affecting forecast inaccuracy (appendix tables A.1 for corn and A.2 for soybeans). Like the anchoring bias, we select the two models with the lowest AIC and BIC for volatility reaction tests and present the results in appendix tables A.3 and A.4 for corn and soybeans, respectively.[8]

For corn, we find that volatility reacts to both predicted and unpredicted absolute surprise components when the decomposition is performed ex post. However, once we account for the latest information set that market participants could use to adjust for the attribution bias, we find that volatility only reacts to the unpredicted absolute surprise obtained from the forecast inaccuracy model with the lowest AIC and BIC. For soybeans, though, volatility response to the predicted absolute surprises, both ex-post and ex-ante, is statistically significant. These findings are opposite to the case with anchoring bias and suggest that while the corn futures market is informationally efficient, the soybean futures market exhibits informational inefficiency in the sense of attribution bias.

Which surprise measure better explains price reaction?

The findings presented in this study demonstrate that the chosen surprise variable, especially when there is any form of bias in market forecasts, affects the measurement of price reactions, thereby influencing the indirect welfare effects of public information. Furthermore, the median

[8] For comparison purposes, we show in column (1) of both tables the results obtained from the ordinary least squares (OLS) regression of volatility (i.e., absolute futures price changes) on the absolute surprise variable taken as the median of firm-level absolute surprises within a year. Column (2) of both tables uses firm-level absolute surprises in a panel regression with firm fixed effects. Columns (3) and (4) decompose the absolute surprise measure into expected and unexpected components by using the predicted absolute residuals from the inaccuracy regression with the two lowest AIC and BIC (columns (1) and (2) in table A.1 for corn, and columns (6) and (2) in table A.2 for soybeans). Columns (3') and (4') present the results obtained with ex-ante decomposed absolute surprise measures in the GMM framework of Campbell and Sharpe (2009).

of industry expectations, typically used to calculate market surprise, ignores the dispersion among analysts' forecasts that may help explain market reaction (Fernandez-Perez et al. 2019). We propose several alternative measures of market surprise that include variability and biases in firm-level forecasts and assess which surprise measure better explains the variation in prices around USDA report releases.

In the simplest form, we use the median of industry forecasts rather than firm-level forecasts while creating the surprise variable (i.e., forecast error) for each year, $SURP_t = 100 \times \frac{Actual_t - median(Forecast_{it})}{Actual_t}$. We then normalize this surprise measure with (a) the standard deviation of firm-level forecasts within a year, (b) the standard deviation of the median surprise across the sample period, (c) the median absolute deviation (MAD) of firm-level forecasts from the consensus, which is taken as the median of forecasts across all firms within a year: $MAD_t = median\left[\left|100 \times \frac{Forecast_{it} - median(Forecast_{it})}{median(Forecast_{it})}\right|\right]$; and (d) the interquartile range (IQR), the difference between the third and first quartile, of firm-level forecasts divided by the median forecast: $IQR_t = Q3\left(100 \times \frac{Forecast_{it})}{median(Forecast_{it})}\right) - Q1\left(100 \times \frac{Forecast_{it})}{median(Forecast_{it})}\right)$. Similar to Fernandez-Perez et al. (2019), we further run regressions with the surprise measure, MAD, and their interaction, as well as regressions with the surprise measure, IQR, and their interaction. The adjusted R-squared values from these regressions estimated via OLS are presented in column (1) of table 5. Column (2) repeats the same analysis by using the firm-level surprises in a panel regression framework. Finally, columns (3) and (4) utilize the firm-level ex-ante decomposed surprises obtained from the anchoring bias models with the minimum and the second-smallest AIC and BIC (corresponding to columns (3') and (4') in tables 3-4). We indicate the models with the largest adjusted R-squared values in bold font in table 5.

For corn, the OLS regression with the surprise variable, created using the median of all firm-level forecasts, and the interquartile range of firm-level forecasts yields the largest explanatory power of price changes around the USDA's August Crop Production report releases, with an adjusted R-squared value of 48%. When compared to the other rows in column (1), it is seen that the largest gain in explanatory power is achieved against the model with the median surprise standardized by the IQR. The smallest gain is over the model with the median surprise and MAD (adjusted R-squared value of 46.6%). The models using the surprises decomposed using the anchoring regression results do not offer higher explanatory power. This might not be surprising, as we demonstrated in table 3 that financial market participants are either unaware of the anchoring bias or fail to account for it in their trading decisions. Among the panel regression models presented in columns (2)-(4), the largest adjusted R-squared is obtained with the model that uses firm-level surprises, IQR, and their interaction.

For soybeans, we find that the model with the decomposed surprises (obtained from the anchoring regression with the second-smallest AIC and BIC), MAD, and their interaction yields the largest explanatory power, with an adjusted R-squared value of 38.6%. While the explanatory power is not substantially different, and sometimes lower, relative to the models in column (3) with decomposed surprises obtained from a different anchoring regression, there is significant improvement over the models using the median surprise presented in column (1). This is not surprising, as we showed in table 4 that financial market participants are aware of the anchoring bias in the industry's soybean production forecasts and make adjustments in their trading decisions. Among the OLS regression models presented in column (1), the largest adjusted R-squared value is obtained with the model that uses the median surprise standardized by the IQR.

Conclusions

Market surprises—differences between USDA crop production estimates and industry forecasts—are widely used to evaluate the impact of USDA reports. While researchers often rely on the median of individual forecasts to represent market expectations, this approach overlooks both the heterogeneity among forecasters and the presence of cognitive biases. Forecast dispersion, for instance, has been shown to influence bid-ask spreads in commodity futures markets during report releases (Fernandez-Perez et al. 2019), suggesting that forecaster diversity carries meaningful information. Moreover, studies such as Karali, Irwin, and Isengildina-Massa (2020) demonstrate that median-based surprises can suffer from attenuation bias, leading to understated price reactions. More recent findings also document cognitive biases—such as anchoring and attribution—in crop production forecasts (Karali, Isengildina-Massa, and Irwin 2025), making parts of the market surprise predictable and potentially exploitable, with implications for market efficiency and welfare analysis.

In this study, we examine the informational efficiency of futures markets by testing whether prices and volatility respond to the predictable component of USDA news driven by anchoring and attribution biases. We find that market participants fail to correct for the reverse anchoring bias in corn forecasts, as evidenced by significant price reactions to the predicted component of the surprise. In contrast, soybean markets appear to adjust for anchoring bias, with prices reacting only to the unpredicted—or true—surprise. These results indicate informational inefficiency in corn futures but efficiency in soybean futures regarding anchoring bias. For attribution bias, the pattern reverses: corn volatility responds only to the unpredicted component, suggesting bias awareness, while soybean volatility reacts to both predicted and unpredicted components, indicating a failure to account for attribution bias.

We further explore which specification of the market surprise measure best captures price variation on days of USDA report releases. Our analysis shows that incorporating the dispersion of firm-level forecasts, such as the interquartile range or median absolute deviation, improves model fit, even when the surprise itself is calculated using the median forecast rather than firm-level data. This suggests that forecast dispersion captures important dimensions of uncertainty and disagreement among forecasters, both of which play a critical role in shaping market reactions. These findings underscore the importance of accounting for heterogeneity among forecasters when assessing the price effects of USDA reports and contribute to a more nuanced understanding of how information is processed in futures markets.

References

Adjemian, M.K. (2012). "Quantifying the WASDE Announcement Effect." *American Journal of Agricultural Economics, 94*(1), 238–257. https://doi.org/10.1093/ajae/aar131

Adjemian, M. K., & Irwin, S. H. (2020). The Market Response to Government Crop News under Different Release Regimes. *Journal of Commodity Markets, 19*:100110. https://doi.org/10.1016/j.jcomm.2019.100110

Campbell, S. D., & Sharpe, S. A. (2009). Anchoring Bias in Consensus Forecasts and Its Effect on Market Prices. *Journal of Financial and Quantitative Analysis, 44*(2), 369–390. https://doi.org/10.1017/S0022109009090127

Colling, P. L., Irwin, S. H., & Zulauf, C. R. (1996). Reaction of Wheat, Corn, and Soybean Futures Prices to USDA 'Export Inspections' Reports. *Review of Agricultural Economics, 18*(1), 127–138. https://doi.org/10.2307/1349672

Dorfman, J. H., & Karali, B. (2015). A Nonparametric Search for Information Effects from USDA Reports. *Journal of Agricultural and Resource Economics, 40*(1), 124–143. http://dx.doi.org/10.22004/ag.econ.197380

Fernandez-Perez, A., Frijns, B., Indriawan, I., & Tourani-Rad, A. (2019). Surprise and Dispersion: Informational Impact of USDA Announcements. *Agricultural Economics, 50*(1), 113–126. https://doi.org/10.1111/agec.12470

Fortenbery, T. R., & Sumner, D. A. (1993). The Effects of USDA Reports in Futures and Options Markets. *Journal of Futures Markets, 13*(2), 157–173. https://doi.org/10.1002/fut.3990130204

Garcia, P., Irwin, S. H., Leuthold, R. M., & Yang, L. (1997). The Value of Public Information in Commodity Futures Markets. *Journal of Economic Behavior and Organization, 32*(4), 559–570. https://doi.org/10.1016/S0167-2681(97)00013-9

Good, D. L., & Irwin, S. H. (2006). *Understanding USDA Corn and Soybean Production Forecasts: Methods, Performance and Market Impacts over 1970–2005*. AgMAS Project Research Report No. 2006-01. http://dx.doi.org/10.2139/ssrn.897279

Hilary, G., & Menzly, L. (2006). Does Past Success Lead Analysts to Become Overconfident? *Management Science, 52*(4), 489–500. https://doi.org/10.1287/mnsc.1050.0485

Hong, H., Kubik, J. D., & Solomon, A. (2000). Security Analysts' Career Concerns and Herding of Earnings Forecasts. *RAND Journal of Economics, 31*(1), 121–144. https://doi.org/10.2307/2601032

Isengildina-Massa, O., Karali, B., & Irwin, S. H. (2017). Do Markets Correct for Smoothing in USDA Crop Production Forecasts? Evidence from Private Analysts and Futures Prices. *Applied Economic Perspectives and Policy, 39*(4), 559–583. https://doi.org/10.1093/aepp/ppx032

Isengildina-Massa, O., Karali, B., & Irwin, S. H. (2024). What Do We Know about the Value and Market Impact of USDA Reports? *Applied Economic Perspectives and Policy, 46*(2), 698–736. https://doi.org/10.1002/aepp.13409

Karali, B. (2012). Do USDA Announcements Affect Comovements Across Commodity Futures Returns? *Journal of Agricultural and Resource Economics, 37*(1), 77–97. http://dx.doi.org/10.22004/ag.econ.122315

Karali, B., Irwin, S. H., & Isengildina-Massa, O. (2020). Supply Fundamentals and Grain Futures Price Movements. *American Journal of Agricultural Economics, 102*(2), 548–568. https://doi.org/10.1002/ajae.12012

Karali, B., Isengildina-Massa, O., & Irwin, S. H. (2025). Is Being Bold Better? Industry Expectations of USDA Corn and Soybean Production Estimates. *Agricultural Economics*. https://doi.org/10.1111/agec.70032

Karali, B., Isengildina-Massa, O., Irwin, S. H., Adjemian, M. K., & Johansson, R. (2019). Are USDA Reports Still News to Changing Crop Markets? *Food Policy, 84*, 66–76. https://doi.org/10.1016/j.foodpol.2019.02.005

McKenzie, A. M. (2008). Pre-Harvest Price Expectations for Corn: The Information Content of USDA Reports and New Crop Futures. *American Journal of Agricultural Economics, 90*(2), 351–366. https://doi.org/10.1111/j.1467-8276.2007.01117.x

McKenzie, A. M., & Ke, Y. (2022). How Do USDA Announcements Affect International Commodity Prices? *Journal of Commodity Markets, 28*:100239. https://doi.org/10.1016/j.jcomm.2021.100239

Milonas, N. T. (1987). The Effects of USDA Crop Announcements on Commodity Prices. *Journal of Futures Markets, 7*(5), 571–589. https://doi.org/10.1002/fut.3990070509

Sumner, D. A., & Mueller, R. A. E. (1989). Are Harvest Forecasts News? USDA Announcements and Futures Market Reactions. *American Journal of Agricultural Economics, 71*(1), 1–8. https://doi.org/10.2307/1241769

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science, 185*(4157), 1124–1131.

Ying, J., Chen, Y., & Dorfman, J. H. (2019). Flexible Tests for USDA Report Announcement Effects in Futures Markets." *American Journal of Agricultural Economics, 101*(4), 1228–1246. https://doi.org/10.1093/ajae/aaz013

Table 1. Anchoring Bias, Corn

	(1) Forecast Error	(2) Forecast Error	(3) Forecast Error	(4) Forecast Error
Deviation from				
Anchor 1	-0.078***			
	(0.008)			
Anchor 2		-0.158***		
		(0.015)		
Anchor 3			-0.096***	
			(0.009)	
Anchor 4				-0.137***
				(0.019)
Constant	0.898***	1.347***	0.876***	1.240***
	(0.025)	(0.060)	(0.019)	(0.067)
Firm effects	Yes	Yes	Yes	Yes
Observations	459	427	443	410
No. of groups	24	24	24	24
Loglikelihood	-1,046.195	-949.428	-1,013.202	-940.089
AIC	2,094.391	1,900.856	2,028.403	1,882.178
BIC	2,098.520	1,904.913	2,032.497	1,886.195

Notes: Estimation results of equation (1) with different anchoring measures are presented. Robust standard errors are in parentheses. Forecast error is defined in percentage terms as $FE_{it} = 100\times[(Actual_t - Forecast_{it}) / Actual_t]$ and deviation from anchor in percentage terms as $100\times[(Forecast_{it} - Anchor_t) / Anchor_t]$. Anchor 1_t = Final $Production_{t-1}$; Anchor 2_t = (Final $Production_{t-1}$+Final $Production_{t-2}$+Final $Production_{t-3}$)/3; Anchor 3_t = $Actual_{t-1}$; Anchor 4_t = ($Actual_{t-1}$+$Actual_{t-2}$+$Actual_{t-3}$)/3. Constant refers to the average of firm fixed effects. AIC refers to Akaike information criteria, and BIC is Bayesian information criteria. The dark (light) gray shaded column represents the model with the (second) smallest AIC and BIC. The asterisks, ***, **, and *, represent statistical significance at the 1%, 5%, and 10%, respectively.

Table 2. Anchoring Bias, Soybeans

	(1) Forecast Error	(2) Forecast Error	(3) Forecast Error	(4) Forecast Error
Deviation from				
Anchor 1	-0.078***			
	(0.008)			
Anchor 2		-0.158***		
		(0.015)		
Anchor 3			-0.096***	
			(0.009)	
Anchor 4				-0.137***
				(0.019)
Constant	0.898***	1.347***	0.876***	1.240***
	(0.025)	(0.060)	(0.019)	(0.067)
Firm effects	Yes	Yes	Yes	Yes
Observations	459	427	443	410
No. of groups	24	24	24	24
Loglikelihood	-1,046.195	-949.428	-1,013.202	-940.089
AIC	2,094.391	1,900.856	2,028.403	1,882.178
BIC	2,098.520	1,904.913	2,032.497	1,886.195

Notes: Estimation results of equation (1) with different anchoring measures are presented. Robust standard errors are in parentheses. Forecast error is defined in percentage terms as $FE_{it} = 100\times[(Actual_t - Forecast_{it}) / Actual_t]$ and deviation from anchor in percentage terms as $100\times[(Forecast_{it} - Anchor_t) / Anchor_t]$. Anchor 1_t = Final $Production_{t-1}$; Anchor 2_t = (Final $Production_{t-1}$+Final $Production_{t-2}$+Final $Production_{t-3}$)/3; Anchor 3_t = $Actual_{t-1}$; Anchor 4_t = ($Actual_{t-1}$+$Actual_{t-2}$+$Actual_{t-3}$)/3. Constant refers to the average of firm fixed effects. AIC refers to Akaike information criteria, and BIC is Bayesian information criteria. The dark (light) gray shaded column represents the model with the (second) smallest AIC and BIC. The asterisks, ***, **, and *, represent statistical significance at the 1%, 5%, and 10%, respectively.

Table 3. Price Reaction Tests with Anchoring Bias, Corn

	(1)	(2)	(3) Ex-post [Table 1(4)]	(3') Ex-ante [Table 1(4)]	(4) Ex-post [Table 1(2)]	(4') Ex-ante [Table 1(2)]
	Futures Return	Futures Return	Futures Return	Futures Return	Futures Return	Futures Return
Median Surprise	-0.999***					
	(0.202)					
Observed Surprise		-0.725***				
		(0.041)				
Predicted Surprise			-1.051***	-0.874***	-1.095***	-1.294***
			(0.101)	(0.133)	(0.081)	(0.119)
Unpredicted Surprise			-0.655***	-0.683***	-0.583***	-0.590***
			(0.061)	(0.067)	(0.063)	(0.072)
Constant	0.316	0.178***	0.543***	0.338**	0.529***	0.455***
	(0.429)	(0.027)	(0.127)	(0.156)	(0.113)	(0.144)
Symmetric Price Reaction (χ^2)			10.76***	1.80	23.80***	22.23***
			[0.00]	[0.18]	[0.00]	[0.00]
Observations	30	459	410	335	427	342
No. of groups		24				
Firm effects		Yes				

Notes: Estimation results of equation (3) are presented with robust standard errors in parentheses. Observed surprise (i.e., forecast error) is defined in percentage terms as $SURP_{it} = 100\times[(Actual_t - Forecast_{it}) / Actual_t]$, where $Forecast_{it}$ is the forecast made by firm i. Median surprise is the median of surprises across firms within a year, defined as $MED(SURP_t) = median(SURP_{it})$. Predicted and unpredicted surprises are, respectively, the predicted dependent variables and predicted residuals from model (4) in table 1 with the minimum AIC and BIC and from model (2) with the second smallest AIC and BIC. To account for the increased sampling variability induced by using generated regressors, models (3)-(4') are estimated via the generalized method of moments (GMM). Symmetric price reaction (Chi-squared test, with p-values in brackets) tests the equality of the coefficient estimates for predicted and unpredicted surprises. The asterisks, ***, **, and *, represent statistical significance at the 1%, 5%, and 10%, respectively.

Table 4. Price Reaction Tests with Anchoring Bias, Soybeans

	(1)	(2)	(3) Ex-post [Table 2(4)]	(3') Ex-ante [Table 2(4)]	(4) Ex-post [Table 2(2)]	(4') Ex-ante [Table 2(2)]
	Futures Return	Futures Return	Futures Return	Futures Return	Futures Return	Futures Return
Median Surprise	-0.555***					
	(0.189)					
Observed Surprise		-0.428***				
		(0.039)				
Predicted Surprise			-0.751***	0.146	-0.282*	0.131
			(0.134)	(0.128)	(0.163)	(0.130)
Unpredicted Surprise			-0.415***	-0.503***	-0.443***	-0.486***
			(0.044)	(0.042)	(0.043)	(0.041)
Constant	0.160	-0.074***	0.065	0.001	-0.088	-0.141
	(0.462)	(0.017)	(0.129)	(0.126)	(0.126)	(0.123)
Symmetric Price Reaction (χ^2)			6.36***	29.43***	0.91	29.41***
			[0.01]	[0.00]	[0.34]	[0.00]
Observations	30	459	410	335	427	342
No. of groups		24				
Firm effects		Yes				

Notes: Estimation results of equation (3) are presented with robust standard errors in parentheses. Observed surprise (i.e., forecast error) is defined in percentage terms as $SURP_{it} = 100\times[(Actual_t - Forecast_{it}) / Actual_t]$, where $Forecast_{it}$ is the forecast made by firm i. Median surprise is the median of surprises across firms within a year, defined as $MED(SURP_t) = median(SURP_{it})$. Predicted and unpredicted surprises are, respectively, the predicted dependent variables and predicted residuals from model (4) in table 2 with the minimum AIC and BIC and from model (2) with the second smallest AIC and BIC. To account for the increased sampling variability induced by using generated regressors, models (3)-(4') are estimated via the generalized method of moments (GMM). Symmetric price reaction (Chi-squared test, with p-values in brackets) tests the equality of the coefficient estimates for predicted and unpredicted surprises. The asterisks, ***, **, and *, represent statistical significance at the 1%, 5%, and 10%, respectively.

Table 5. Explanatory Power of Surprise Measures

Dep. var.: Futures Return	(1) Median	(2) Firm-level	(3) Ex-ante [Table 3(3')] Firm-level Decomposed	(4) Ex-ante [Table 3(4')] Firm-level Decomposed
		Corn		
Surprise	0.426	0.351	0.317	0.354
Surprise ÷ Std. Dev. (Forecasts)	0.350	0.313	0.296	0.315
Surprise ÷ Std. Dev. (Surprise)	0.426	0.351	0.296	0.315
Surprise ÷ MAD	0.329	0.269	0.195	0.227
Surprise ÷ IQR	0.266	0.205	0.137	0.158
Surprise, MAD	0.466	0.375	0.345	0.362
Surprise, MAD, Surprise × MAD	0.446	0.398	0.345	0.362
Surprise, IQR	**0.480**	0.402	0.381	0.386
Surprise, IQR, Surprise × IQR	0.460	0.423	0.381	0.386
		Soybeans		
Surprise	0.221	0.214	0.376	0.380
Surprise ÷ Std. Dev. (Forecasts)	0.222	0.212	0.374	0.371
Surprise ÷ Std. Dev. (Surprise)	0.221	0.214	0.374	0.371
Surprise ÷ MAD	0.157	0.069	0.333	0.327
Surprise ÷ IQR	0.232	0.216	0.360	0.344
Surprise, MAD	0.215	0.219	0.383	**0.386**
Surprise, MAD, Surprise × MAD	0.185	0.217	0.383	**0.386**
Surprise, IQR	0.192	0.212	0.374	0.378
Surprise, IQR, Surprise × IQR	0.194	0.227	0.374	0.378

Notes: The table presents adjusted R-squared values from the regression of daily futures returns calculated as $\Delta P_t = 100 \times (\ln P_t - \ln P_{t-1})$, where P_t is the price of the new crop futures contract, on various surprise measures. MAD refers to the median absolute deviation, and IQR refers to the interquartile range. The dark and light gray shaded column uses the predicted and unpredicted surprises obtained from model (4) in tables 1 and 2 with the minimum AIC and BIC and from model (2) with the second smallest AIC and BIC, respectively. The maximum adjusted R-squared values across the models for each commodity are indicated with a bold font.

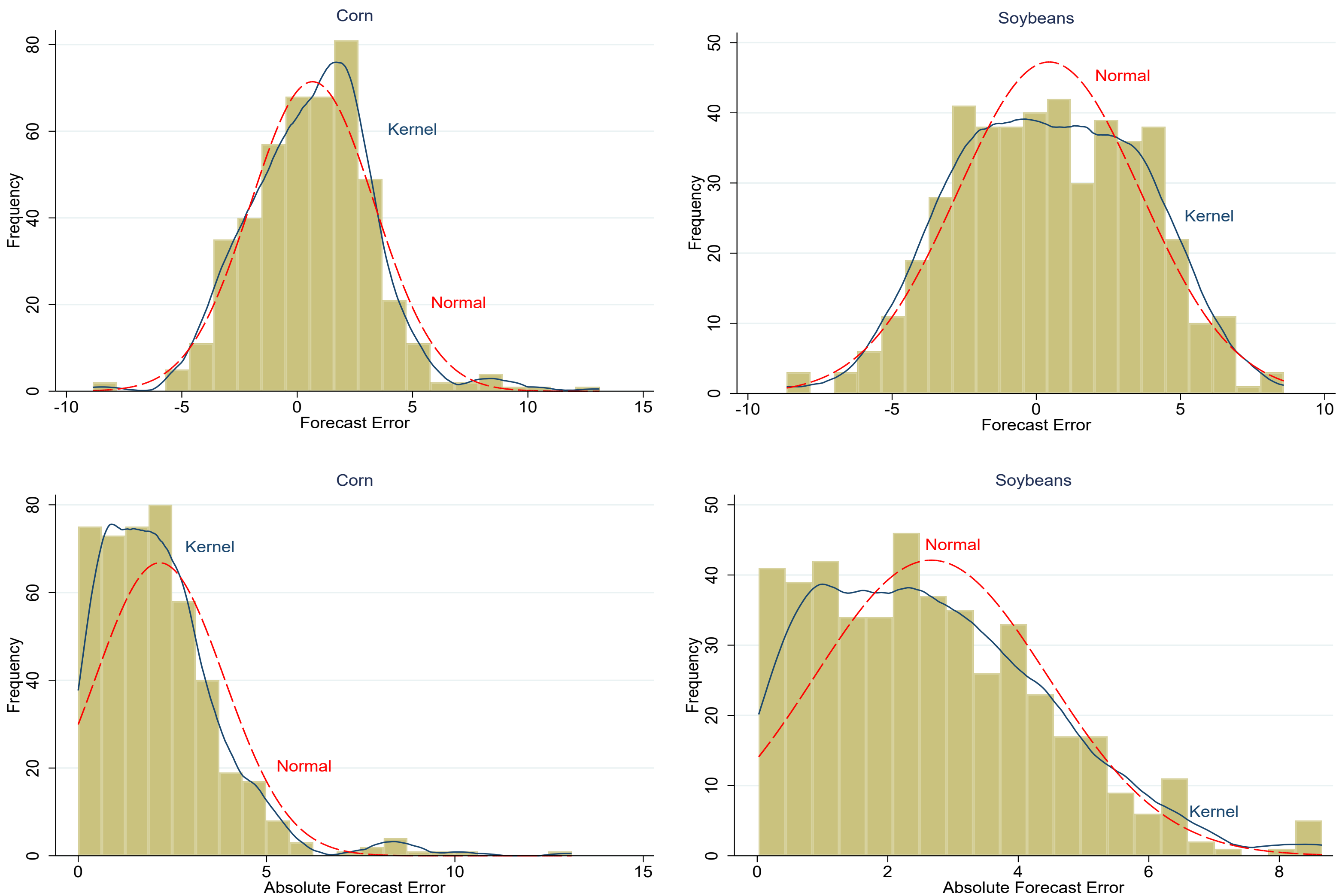

Figure 1. Empirical probability distributions of forecast errors

Notes: Forecast error is defined in percentage terms as $FE_{it} = 100\times[(Actual_t - Forecast_{it}) / Actual_t]$, and absolute forecast error is defined as $|FE_{it}| = |100\times[(Actual_t - Forecast_{it}) / Actual_t]|$.

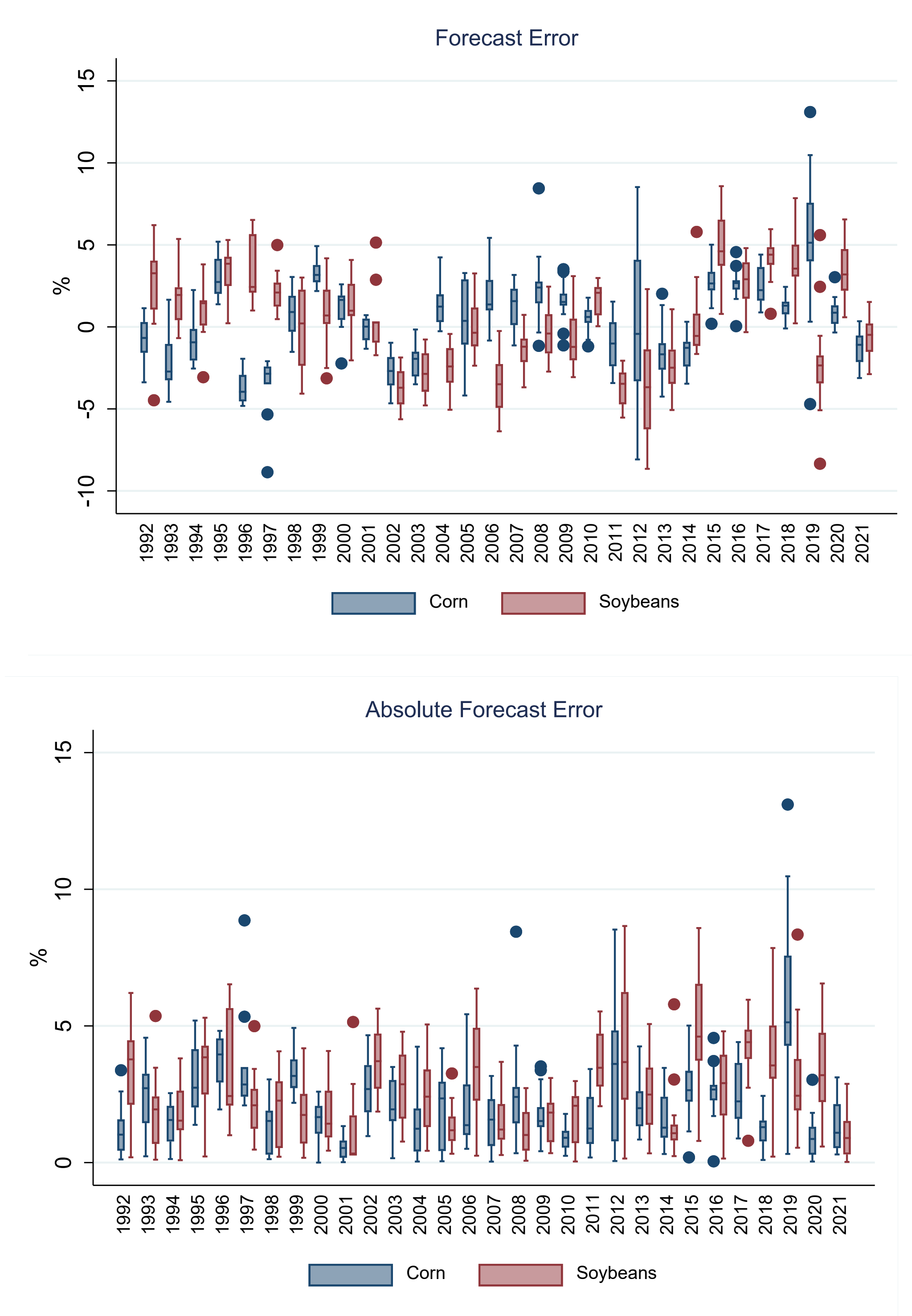

Figure 2. Forecast errors across firms over time

Notes: Forecast error is defined in percentage terms as $FE_{it} = 100\times[(Actual_t - Forecast_{it}) / Actual_t]$, and absolute forecast error is defined as $|FE_{it}| = |100\times[(Actual_t - Forecast_{it}) / Actual_t]|$.

Appendix: Attribution Bias and Volatility Reaction Tests

In this appendix, we explain the empirical models used to verify the existence of attribution bias based on the econometric framework in Karali, Isengildina-Massa, and Irwin (2025). After selecting the two models with the lowest Akaike information criteria (AIC) and Bayesian information criteria (BIC), we perform volatility reaction tests to the predicted component of the surprise induced by cognitive biases.

Attribution bias is defined as overconfidence in analysts' own forecasting skills resulting from having success in previous periods, making them deviate more from the consensus, which is taken as the median of others' forecasts (Hong, Kubik, and Solomon 2000; Hilary and Menzly 2006). Karali, Isengildina-Massa, and Irwin (2025) investigated the existence of an attribution bias using two different approaches. In one approach, they model the absolute value of forecast deviations from the consensus, and in the other, they model the absolute value of forecast errors. Since the absolute value of forecast errors corresponds to the absolute value of market surprises, the topic of interest in our study, we estimate the following regression equation that incorporates both firm and year fixed effects:

(A.1) $$|FE_{it}| = \pi_i + \gamma_t + \phi\left|FE_{i,t-1}\right| + \beta\ Boldness\ Score_{it} + \delta\ Experience_{it} + \lambda\ Freq_{it} + \epsilon_{it}.$$

The dependent variable, $|FE_{it}|$, is the absolute value of the forecast error defined as $FE_{it} = 100 \times \frac{(Actual_t - Forecast_{it})}{Actual_t}$, where $Actual_t$ is the actual value in year *t* and $Forecast_{it}$ is the forecast made by firm *i* in year *t* (see equation (2) in the paper). The parameters π_i and γ_t represent the firm and time effects, respectively. A bold forecast is defined as one that significantly deviates from the consensus, where consensus is taken as the median (or mean) of the other analysts' forecasts. We calculate the boldness score of each firm following Hong, Kubik, and Solomon (2000) as:

$$\text{(A.2)} \quad Score_{it} = 100 - \left(\frac{Rank_{it} - 1}{Number\ of\ firms_t - 1}\right) \times 100.$$

We assign $Rank_{it}$ to each firm for each year separately by ranking the absolute forecast deviation, $|FDEV_{it}|$, from the largest to the smallest value, with the absolute forecast deviation in percentage terms is calculated as:

$$\text{(A.3)} \quad |FDEV_{it}| = \left|100 \times \frac{\left(Forecast_{it} - Forecast_{-i,t}\right)}{Forecast_{-i,t}}\right|.$$

The variable $Forecast_{-i,t}$ is the median of the forecasts made by other firms (excluding firm *i*) in year *t*, which implicitly assumes that each firm has perfect knowledge about the forecasts made by other firms. With this setup, the firm with the largest deviation in magnitude takes the rank of one since a larger forecast deviation from the consensus indicates more boldness; the firm with the second-largest forecast deviation has the rank of two, and so on. In the case of equal absolute forecast deviations for two or more firms, we use the mid-point of the rank for those firms; therefore, the boldness rank is not necessarily an integer. As a result, the firm with the largest forecast deviation receives a score of 100, while the firm with the smallest forecast deviation receives a score of zero. The median score in each year is 50 by construction. As in Karali, Isengildina-Massa, and Irwin (2005), we proxy a firm's experience by the running total of years the company provides forecasts, with the experience variable taking the value of one in the first year the company provides a forecast, and increasing by one in the year when that company makes a forecast. We create the variable $Freq_{it}$ to serve as a proxy for the frequency of superior forecasts, following Karali, Isengildina-Massa, and Irwin (2025), who used a similar approach to Hilary and Menzly (2006). Specifically, we count the number of times a firm's absolute forecast error was below the median of other firms' forecast errors (i.e., superior forecast) in the last three periods, $Freq_{it} = Count(|FE_{i,t-j}| < |FE_{-i,t-j}|),\ j = 1, 2, 3$.[9] Positive parameter estimates would indicate that forecast errors increase in magnitude with previous forecast errors, boldness, experience, and prior success.

Similar to the case with anchoring bias, the finding of statistically significant slope estimates in equation (A.1) suggests that the absolute forecast error, which is equivalent to the absolute market surprise, is partly predictable. Thus, we can decompose the absolute surprise measure into expected and unexpected components, where the expected component is simply the predicted dependent variable from equation (A.1.), $|\widehat{FE}_{it}|$, and the unexpected component is the predicted residuals, $|\hat{\epsilon}_{it}| = |FE_{it}| - |\widehat{FE}_{it}|$. We can then test for volatility reactions by estimating the following regression equation:

(A.4) $$|\varDelta P_{it}| = \theta_i + \psi_1 |\widehat{FE}_{it}| + \psi_2\, |\hat{\epsilon}_{it}| + \upsilon_{it},$$

where $|\varDelta P_{it}| = |100 \times (\ln P_t - \ln P_{t-1})|$ is the absolute value of continuously compounded daily return on the commodity futures contract with price P_t, which is a commonly-used volatility measure in the literature, and it is the same for all firms ($|\varDelta P_{it}| = |\varDelta P_t|, \forall i$). If the markets are informationally efficient, then $\psi_1 = 0$, indicating that market participants are aware

[9] We estimate equation (A.1) using different regressors, which are presented in table A.1. The accuracy score is calculated using equation (A.2), where we assign $Rank_{it}$ for each year separately by ranking the absolute forecast error, $|FE_{it}|$, from the smallest to the largest value. Accordingly, the firm with the smallest forecast error receives the rank of one, the firm with the second-smallest forecast error receives the rank of two, and so on. When forecast errors are equal for two or more firms, we use the mid-point of the rank; therefore, the accuracy rank is not necessarily an integer. As a result, the accuracy score of the firm with the smallest forecast error is 100, and that of the firm with the largest forecast error is zero. By definition, the median score in each year is 50. We also create indicator variables for the accuracy and boldness scores and ranks for being in the top 5th percentile and the bottom 5th percentile of their respective distribution.

of the attribution bias in industry forecasts and, therefore, volatility does not respond to the anticipated surprise component. We estimate equation (A.4) using the GMM method in Campbell and Sharpe (2009) to account for the increased sampling variability introduced by using generated regressors.

Table A.1. Determinants of Forecast Inaccuracy, Corn

	(1) Absolute Forecast Error	(2) Absolute Forecast Error	(3) Absolute Forecast Error	(4) Absolute Forecast Error	(5) Absolute Forecast Error	(6) Absolute Forecast Error
Prior Absolute Forecast Error	-0.166***	-0.165***				
	(0.041)	(0.050)				
Prior Accuracy Score			0.007**			
			(0.003)			
Boldness Score	0.012***		0.012***			
	(0.003)		(0.003)			
Prior Accuracy Rank				-0.040**		
				(0.016)		
Boldness Rank				-0.058***		
				(0.018)		
Prior Top 5% Accuracy Score					0.555***	0.536***
					(0.185)	(0.176)
Prior Bottom 5% Accuracy Score						-0.095
						(0.240)
Top 5% Boldness Score		1.172**			1.144**	1.090**
		(0.461)			(0.458)	(0.463)
Bottom 5% Boldness Score						-0.312*
						(0.177)
Experience	0.078*	0.093**	0.069*	0.074*	0.081**	0.080*
	(0.042)	(0.044)	(0.039)	(0.040)	(0.040)	(0.040)
Prior Success Frequency	0.119	0.145*	0.094	0.110	0.210***	0.208***
	(0.074)	(0.085)	(0.091)	(0.091)	(0.076)	(0.074)
Constant	2.307***	2.781***	1.777***	3.233***	2.407***	2.459***
	(0.497)	(0.479)	(0.492)	(0.478)	(0.466)	(0.450)
Firm effects	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	316	316	316	316	316	316
No. of groups	24	24	24	24	24	24
Loglikelihood	-473.653	-477.529	-477.258	-477.100	-479.065	-477.964
AIC	993.305	1,001.057	994.515	1,000.200	1,004.129	1,001.929
BIC	1,079.687	1,087.439	1,080.897	1,086.583	1,090.511	1,088.311

Notes: Estimation results of equation (A.1) are presented along with alternative control variables. Robust standard errors are in parentheses. Forecast error is defined in percentage terms as $FE_{it} = 100\times[(Actual_t - Forecast_{it}) / Actual_t]$, and forecast deviation in percentage terms as $FDEV_{it} = 100\times[(Forecast_{it} - Forecast_{-i,t}) / Forecast_{-i,t}]$, where $Forecast_{-i,t}$ is the median of the forecasts made by other firms. Accuracy and boldness scores are calculated as $Score_{it} = 100 - [(Rank_{it}-1) / (\text{Number of firms}_t-1)]\times 100$, where $Rank_{it}$ is assigned for the accuracy score by ranking the absolute forecast error from the smallest to the largest value, with the smallest forecast error having the rank of one, and for the boldness score by ranking the absolute forecast deviation from the largest to the smallest value, with the largest forecast deviation having the rank of one. Experience is the running total of years firms provided a forecast and prior success frequency is the number of times a firm's absolute forecast error was below the median of other firms' forecasts errors in the last three periods, calculated as $Freq_{it} = \text{Count}(|FE_{i,t-j}| < |FE_{-i,t-j}|)$, j=1, 2, 3. Other variables are indicator variables for forecasts being in the specified percentile of the accuracy or boldness score distributions. Constant refers to the average of firm fixed effects. AIC refers to Akaike information criteria and BIC is Bayesian information criteria. The dark (light) gray shaded column represents the model with the (second) smallest AIC and BIC. The asterisks, ***, **, and *, represent statistical significance at the 1%, 5%, and 10%, respectively.

Table A.2. Determinants of Forecast Inaccuracy, Soybeans

	(1) Absolute Forecast Error	(2) Absolute Forecast Error	(3) Absolute Forecast Error	(4) Absolute Forecast Error	(5) Absolute Forecast Error	(6) Absolute Forecast Error
Prior Absolute Forecast Error	-0.042 (0.050)	-0.040 (0.047)				
Prior Accuracy Score			0.003 (0.003)			
Boldness Score	0.014*** (0.003)		0.014*** (0.003)			
Prior Accuracy Rank				-0.022 (0.018)		
Boldness Rank				-0.074*** (0.016)		
Prior Top 5% Accuracy Score					0.081 (0.312)	0.068 (0.307)
Prior Bottom 5% Accuracy Score						-0.061 (0.275)
Top 5% Boldness Score		1.617*** (0.465)			1.617*** (0.468)	1.577*** (0.487)
Bottom 5% Boldness Score						-0.340* (0.188)
Experience	-0.077 (0.084)	-0.100 (0.080)	-0.075 (0.085)	-0.084 (0.084)	-0.096 (0.080)	-0.093 (0.082)
Prior Success Frequency	0.127 (0.088)	0.134 (0.088)	0.093 (0.089)	0.080 (0.091)	0.161** (0.079)	0.144 (0.093)
Constant	2.448*** (0.814)	2.975*** (0.781)	2.253*** (0.743)	3.688*** (0.651)	2.858*** (0.703)	2.931*** (0.757)
Firm effects	Yes	Yes	Yes	Yes	Yes	Yes
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	316	316	316	316	316	316
No. of groups	24	24	24	24	24	24
Loglikelihood	-507.874	-506.912	-507.304	-509.602	-507.093	-506.172
AIC	1,061.747	1,059.825	1,060.607	1,065.204	1,060.186	1,058.344
BIC	1,148.129	1,146.207	1,146.989	1,151.586	1,146.568	1,144.726

Notes: Estimation results of equation (A.1) are presented along with alternative control variables. Robust standard errors are in parentheses. Forecast error is defined in percentage terms as $FE_{it} = 100\times[(Actual_t - Forecast_{it}) / Actual_t]$, and forecast deviation in percentage terms as $FDEV_{it} = 100\times[(Forecast_{it} - Forecast_{-i,t}) / Forecast_{-i,t}]$, where $Forecast_{-i,t}$ is the median of the forecasts made by other firms. Accuracy and boldness scores are calculated as $Score_{it} = 100 - [(Rank_{it}-1) / (Number\ of\ firms_t-1)]\times 100$, where $Rank_{it}$ is assigned for the accuracy score by ranking the absolute forecast error from the smallest to the largest value, with the smallest forecast error having the rank of one, and for the boldness score by ranking the absolute forecast deviation from the largest to the smallest value, with the largest forecast deviation having the rank of one. Experience is the running total of years firms provided a forecast and prior success frequency is the number of times a firm's absolute forecast error was below the median of other firms' forecasts errors in the last three periods, calculated as $Freq_{it} = Count\ (|FE_{i,t-j}| < |FE_{-i,t-j}|)$, j=1, 2, 3. Other variables are indicator variables for forecasts being in the specified percentile of the accuracy or boldness score distributions. Constant refers to the average of firm fixed effects. AIC refers to Akaike information criteria and BIC is Bayesian information criteria. The dark (light) gray shaded column represents the model with the (second) smallest AIC and BIC. The asterisks, ***, **, and *, represent statistical significance at the 1%, 5%, and 10%, respectively.

Table A.3. Volatility Reaction Tests with Attribution Bias, Corn

	(1)	(2)	(3) Ex-post [Table A.1(1)]	(3') Ex-ante [Table A.1(1)]	(4) Ex-post [Table A.1(3)]	(4') Ex-ante [Table A.1(3)]
	Volatility	Volatility	Volatility	Volatility	Volatility	Volatility
Median Absolute Surprise	0.773***					
	(0.244)					
Observed Absolute Surprise		0.415***				
		(0.043)				
Predicted Absolute Surprise			0.666***	0.097	0.677***	-0.136*
			(0.053)	(0.130)	(0.053)	(0.075)
Unpredicted Absolute Surprise			0.008	0.484***	0.010	0.114*
			(0.095)	(0.070)	(0.098)	(0.067)
Constant	0.988*	1.748***	1.284***	1.948***	1.264***	2.947***
	(0.546)	(0.093)	(0.131)	(0.278)	(0.127)	(0.149)
Symmetric Volatility Reaction (χ^2)			26.14***	5.94***	25.74***	3.14*
			[0.00]	[0.01]	[0.00]	[0.08]
Observations	30	459	316	260	316	260
No. of groups		24				
Firm effects		Yes				

Notes: Estimation results of equation (A.4) are presented with robust standard errors in parentheses. Observed absolute surprise (i.e., absolute forecast error) is defined in percentage terms as $|SURP_{it}| = |\ 100 \times [(Actual_t - Forecast_{it}) / Actual_t]\ |$, where $Forecast_{it}$ is the forecast made by firm i. Median absolute surprise is the median of absolute surprises across firms within a year, defined as $MED(|SURP_t|) = median(\ |SURP_{it}|\ )$. Predicted and unpredicted absolute surprises are, respectively, the predicted dependent variables and predicted residuals from model (1) in table A.1 with the minimum AIC and BIC and from model (3) with the second smallest AIC and BIC. To account for the increased sampling variability induced by using generated regressors, models (3)-(4') are estimated via the generalized method of moments (GMM). Symmetric volatility reaction (Chi-squared test, with p-values in brackets) tests the equality of the coefficient estimates for predicted and unpredicted absolute surprises. The asterisks, ***, **, and *, represent statistical significance at the 1%, 5%, and 10%, respectively.

Table A.4. Volatility Reaction Tests with Attribution Bias, Soybeans

	(1)	(2)	(3) Ex-post [Table A.2(6)]	(3') Ex-ante [Table A.2(6)]	(4) Ex-post [Table A.2(3)]	(4') Ex-ante [Table A.2(3)]
	Volatility	Volatility	Volatility	Volatility	Volatility	Volatility
Median Absolute Surprise	0.671**					
	(0.278)					
Observed Absolute Surprise		0.354***				
		(0.051)				
Predicted Absolute Surprise			0.602***	-0.145***	0.579***	-0.158***
			(0.076)	(0.059)	(0.076)	(0.053)
Unpredicted Absolute Surprise			-0.296***	0.096**	-0.268***	0.101***
			(0.108)	(0.050)	(0.106)	(0.046)
Constant	0.503	1.220***	0.738***	2.320***	0.773***	2.318***
	(0.684)	(0.134)	(0.215)	(0.165)	(0.211)	(0.155)
Symmetric Volatility Reaction (χ^2)			35.58***	5.18**	32.31***	7.06***
			[0.00]	[0.02]	[0.00]	[0.01]
Observations	30	459	316	260	316	260
No. of groups		24				
Firm effects		Yes				

Notes: Estimation results of equation (A.4) are presented with robust standard errors in parentheses. Observed absolute surprise (i.e., absolute forecast error) is defined in percentage terms as $|SURP_{it}| = |\ 100 \times [(Actual_t - Forecast_{it}) / Actual_t]\ |$, where $Forecast_{it}$ is the forecast made by firm i. Median absolute surprise is the median of absolute surprises across firms within a year, defined as $MED(|SURP_t|) = median(\ |SURP_{it}|\ )$. Predicted and unpredicted absolute surprises are, respectively, the predicted dependent variables and predicted residuals from model (1) in table A.2 with the minimum AIC and BIC and from model (3) with the second smallest AIC and BIC. To account for the increased sampling variability induced by using generated regressors, models (3)-(4') are estimated via the generalized method of moments (GMM). Symmetric volatility reaction (Chi-squared test, with p-values in brackets) tests the equality of the coefficient estimates for predicted and unpredicted absolute surprises. The asterisks, ***, **, and *, represent statistical significance at the 1%, 5%, and 10%, respectively.