Does Competition Improve Information Quality: Evidence from the Security Analyst Market

Chuqing Jin

Toulouse School of Economics *

September 2026

Abstract

Experts are often rewarded for relative performance. For security analysts, this turns forecasting stocks into a contest in which analysts compete to be more accurate than their peers. The contest can induce analysts to distort their forecasts to stand out, but can also discipline them against issuing overly optimistic forecasts—a central concern of policymakers in this market. I develop and structurally estimate a forecasting contest model, inferring analysts' incentives from their forecasts using a revealed-preference approach. The model incorporates relative-accuracy and optimism incentives in a unified framework, enabling counterfactual simulations that speak to recent policy debates over optimism and analyst competition. I find that the incentives interact, leading analysts with bullish signals to disproportionately exaggerate them. Under the estimated incentives, the contest nevertheless acts as a "market-based" discipline on optimism, reducing forecast errors overall even as its interaction with optimism amplifies forecast noise. The estimated model implies that aggregate information quality is inverted-U-shaped in the number of competing analysts: additional competition produces more information but also intensifies contest-driven distortions.

*Email: *chuqing.jin@tse-fr.eu*. I would like to thank my advisors Marc Rysman, Jihye Jeon, and Hiroaki Kaido. Thanks to Allan Timmermann, Chiara Margaria, Charles Hodgson (discussant), Daniel Garrett, Gaston Illanes (discussant), Glenn Ellison (discussant), Juan Ortner, Karam Kang, Kevin Pisciotta (discussant), Lucy White, Lulu Wang (discussant), Marina Halac, Mark Egan, Marco Ottaviani, Navin Kartik, Roger Loh, Todd Milbourn, and Zachary Kaplan and seminar participants at Bank of Canada, the Bank of Canada-Queen's Workshop on Financial Intermediation and Regulation, Barcelona School of Economics Summer Forum, Boston University, Carnegie Mellon University, the Center for the Study of Industrial Organization - Toulouse School of Economics Joint Workshop on Industrial Organization, CEPR-JIE Conference on Applied Industrial Organization, Chicago Household Finance Conference, Econometric Society European Winter Meeting, Econometric Society World Congress, the European Association for Research in Industrial Economics, ESSEC Business School, the Financial Management Association Annual Meeting, HEC Paris, the International Industrial Organization Conference, Nanyang Technological University, National University of Singapore, Norwegian School of Economics, Singapore Management University, Toulouse School of Economics, Ohio State University, Queen's University, and University of Oxford for helpful conversations and suggestions at various stages of the paper. All errors are my own. I acknowledge funding from ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program). JEL codes: D80, G20, L10

1 Introduction

The efficiency of many markets depends on information provided by experts, and markets tend to reward the experts who prove most accurate relative to their peers. This reward for *relative accuracy* turns forecasting into a contest between the experts (Ottaviani and Sørensen, 2006), so forecasts become strategic choices rather than statements of honest beliefs. In financial markets, such incentives are particularly prevalent. Security analysts compete to forecast stocks and are rewarded for being *more* accurate than their peers (Hong and Kubik, 2003),[1] but their forecasts are often taken as their honest beliefs, especially by retail investors. If analysts issue distorted forecasts to maximize their own payoffs, these forecasts may mislead investors and lead to bad investment decisions.

This paper studies the effect of competition on information quality in financial markets. I focus on security analysts, who are experts employed by brokerage houses to forecast the stocks they specialize in. Competition here takes the form of a forecasting contest, where analysts compete to issue more accurate forecasts than their rivals. The main challenge in understanding analysts' strategic incentives is that their payoffs are largely unobservable, as brokerages rarely disclose analysts' compensation. To resolve this unobservability, existing studies often rely on career outcomes as proxies (Hong and Kubik, 2003). This paper instead takes a revealed-preference approach: I build and estimate a structural model of analysts' forecasting contest to infer their payoffs using forecast data. The estimated model allows me to simulate counterfactual policies to quantify their impact on information quality.

I use the model to study counterfactuals that speak to policymakers' long-standing concern about analyst conflicts of interest. In particular, brokerages may reward analysts for issuing overly optimistic forecasts, which help the brokerage generate revenue from trading commissions and investment-banking business. But regulating this optimism might have an adverse effect on competition, and hence information quality. Specifically, regulators often attempt to insulate analyst research from such conflicting revenue sources, which in the past has resulted in lower budgets for analyst research and reduced analyst coverage. In the United States (US), the Global Research Analyst Settlement (the Settlement) separated analysts' compensation from investment-banking revenues at major bro-

[1] Analysts are also routinely ranked in "best-analyst" awards and winners of these awards could receive a huge boost in compensation and career outcomes. Thomson Reuters Analyst Awards explicitly uses earnings forecast accuracy as the ranking criterion, whereas forecast accuracy enters indirectly in other awards. See, for example, `https://www.institutionalinvestor.com/article/2bsw4ehe37jv5y8dp454w/corner-office/the-all-america-research-team-jpmorgan-triumphs-again` and `https://www.thomsonreuters.com/en/press-releases/2018/august/thomson-reuters-analyst-awards-2018-winners-of-us-awards-announced`, accessed August 10, 2026.

kerages in 2003; in Europe, the unbundling rule of the Markets in Financial Instruments Directive II (MiFID II) separated payments for analyst research from trading commissions in 2018.[2]

These policies were enacted to curb optimism, but policymakers remained apprehensive that they could weaken competition and impair information quality. The regulatory pendulum has recently swung back: since 2024, policymakers in the US and Europe have rolled back these policies to varying degrees, largely to promote competition.[3] This policy reversal highlights persistent uncertainty about the interaction between optimism and competition.

I contribute to the ongoing debate by bringing both forces into a unified framework to study their interaction. My model allows analysts' payoffs to depend on relative accuracy, optimism, and absolute accuracy. Analysts compete in a contest to forecast the earnings of a security in a given year. They receive private signals about the earnings and choose their forecasts simultaneously. Once the earnings are realized, analysts receive payoffs based on their own forecasts, their rivals' forecasts, and the realized earnings.

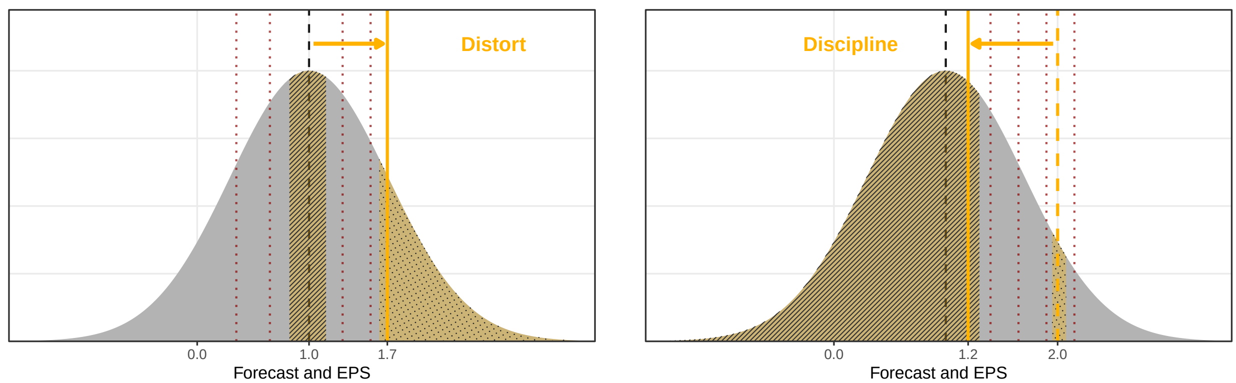
To illustrate the main intuition of the model, consider a simple thought experiment. Suppose an analyst believes, based on her private signal, that the expected earnings will be \$1. Figure 1a illustrates her trade-offs when choosing a forecast. Assume for this illustration only that she knows her rivals' forecasts, which are distributed around her posterior mean. If she only cares about absolute accuracy, she would report her honest belief—the posterior mean of \$1—as her forecast. However, if she competes in a winner-takes-all forecasting contest and is rewarded for being the most accurate analyst, she may distort her forecast away from her honest belief—say to \$1.7. This differentiates her from her rivals and raises her chance of being the most accurate: if the earnings turn out to be sufficiently high, she would be the only winner; in contrast, her honest forecast would only be the most accurate over a narrow range of earnings realizations. This shows the potential *distortionary* effect of rewarding relative accuracy. The magnitude of this distortion also depends on the level of competition: when the analyst faces more rivals, it becomes

[2]See https://www.sec.gov/news/press/2003-54.htm and https://www.esma.europa.eu/publications-and-data/interactive-single-rulebook/mifid-ii/article-24-general-principles-and, accessed August 3, 2026.

[3]The Securities and Exchange Commission (SEC) removed the remaining undertakings of the Settlement in 2025 partly due to its "chilling effect on research coverage". The United Kingdom (UK) reversed the MiFID II unbundling rule in 2024 and the European Union (EU) adopted a similar reversal in its 2024 Listing Act, both using promoting competition as a main supporting argument. See https://www.sec.gov/newsroom/speeches-statements/uyeda-statement-global-research-analyst-settlement-120525, https://www.fca.org.uk/publications/policy-statements/ps24-9-payment-for-optionality-investment-research, and https://eur-lex.europa.eu/eli/dir/2024/2811/oj/eng, accessed August 4, 2026.

Figure 1: Distortionary and Disciplinary Effects of Relative Accuracy

(a) Distortionary effect (b) Disciplinary effect



Note: Figure illustrates an analyst's trade-offs when choosing a forecast given different payoffs. The red dotted vertical lines represent her rivals' forecasts, which are assumed to be known to the analyst for this illustration only. The gray area represents the analyst's posterior distribution of the earnings. The black dashed line marks the analyst's posterior mean, i.e., her *honest forecast*. The yellow pattern-shaded areas represent the probabilities of being the most accurate for different forecasts. Numerical values are illustrative. In Panel a, the yellow solid line is the analyst's *distorted forecast*; the yellow arrow marks the direction of the distortion. The striped (dotted) pattern on the density plot marks the earnings realizations for which her *honest (distorted) forecast* would be the most accurate. In Panel b, the yellow dashed line is the *optimism-only* forecast induced by the optimism incentive without a reward for relative accuracy. The yellow solid line is the *disciplined forecast* when the analyst also competes on relative accuracy; the yellow arrow represents the direction of the discipline. The striped (dotted) pattern marks the earnings realizations for which the *disciplined (optimism-only) forecast* would be the most accurate.

harder for her to differentiate herself, so she may respond by distorting more (illustrated in Appendix Figure A.1).

Now suppose analysts also have an incentive to issue optimistic forecasts. Because her rivals are likely optimistic as well, the analyst can raise her chance of being the most accurate by issuing a less optimistic forecast (Figure 1b). Here rewarding relative accuracy moves her forecast towards her honest belief and has a *disciplinary* effect on optimism. When a policy such as the Settlement or MiFID II changes both the optimism incentive and the level of competition, its net effect depends on the relative magnitude of the distortionary and disciplinary forces and becomes an empirical question.

My structural model allows relative accuracy and optimism to enter flexibly in analysts' payoff. Importantly, I do not impose relative accuracy or optimism incentives on analysts, but instead infer what these parameters must be to rationalize their forecasts. I estimate the model with indirect inference, using data on earnings forecasts and realized earnings from the Institutional Brokers' Estimate System (IBES) between 1984 and 2016. The realized earnings allow me to evaluate analysts' forecast quality ex post. The indirect inference approach searches for model parameters that allow the model to reproduce key empirical patterns.

The main challenge in estimating the model is to jointly identify analysts' payoff and signal distribution. For example, when we observe an analyst making a high earnings forecast, it is hard to distinguish whether it is due to a high private signal or a payoff that incentivizes exaggeration or optimism. To address this challenge, I exploit variation in forecast strategy with the level of competition to separately identify the payoff. The intuition follows from the thought experiment in Figure 1: in a winner-takes-all contest, analysts distort their forecasts to differentiate themselves; the more rivals they face, the harder it is to be the most accurate, so they distort more, resulting in higher forecast errors. Therefore, if analysts face a large reward for the most accurate, we expect their forecast errors to increase with the level of competition. Using this variation, I establish sufficient conditions under which my model is identified.

My structural estimates show that analysts receive a significant reward both for issuing the most accurate forecast and for optimism. This implies that analysts are not only incentivized to issue optimistic forecasts, but they also strategically choose their forecasts to increase their chance of winning a winner-takes-all forecasting contest. The magnitudes of these two forces relative to the baseline reward for absolute accuracy determine the equilibrium forecasting strategy. The optimism incentive is strong, which leads nearly all analysts to bias their forecasts upwards. Nevertheless, the reward for the most accurate gives them an incentive to differentiate themselves, so analysts with positive signals distort their forecasts upwards more than analysts with negative signals. In contrast, if the optimism incentive were shut down, analysts would only modestly distort their forecasts towards their private signals: those with positive signals distort their forecasts upwards, and those with negative signals, downwards.

The estimated model provides a unified framework to study how optimism and competition interact in this market. Though it does not microfound why optimism regulations reduce competition, it can simulate policies that affect both optimism and competition in flexible ways. To this end, I simulate counterfactuals that eliminate rewards for relative accuracy and optimism separately or simultaneously, and quantify their impact on information quality at different levels of competition. I consider the information quality both of any individual analyst and the aggregate market as reflected in consensus forecasts. The consensus is defined as the mean forecast in each security-year and is the most frequently quoted forecast in financial news outlets.

I find that the optimism incentive not only makes forecasts more biased but also noisier due to an interaction with rewards for relative accuracy. This is because optimism has an asymmetric effect on analysts with different signals, as described above in the equilibrium strategy. For analysts with positive signals, the incentive to be the most accurate and optimistic both lead them to distort their forecasts upwards, resulting in even larger distortions, whereas for analysts with negative signals, these two incentives go in opposite directions. Therefore, when analysts strategically interact with each other, optimism magnifies the noise in the market. What if we stop comparing analysts against each other and remove the winner-takes-all contest? It turns out the contest is still net positive. Once the reward for the most accurate is eliminated, the strong optimism incentive leads analysts to report extremely high forecasts regardless of their private signals. The *disciplinary* effect dominates in the current market: the reward for relative accuracy reduces the individual and consensus forecast errors by 34% and 61%, respectively. However, as the contest still gives analysts an incentive to differentiate, this improvement in information accuracy comes at the cost of forecasts being 7% noisier.

While the contest is net positive overall, more competition is not strictly better. Competition has an inverted-U-shaped effect on the quality of the consensus forecast. More analysts bring more information to the market, which improves consensus quality initially. However, as the level of competition increases further, it becomes more difficult for analysts to differentiate themselves and win the forecasting contest, so the distortionary effect intensifies and impairs information quality. This eventually cancels out the positive effect of more information. The inverted-U curve explains in part why additional analyst coverage is often beneficial for under-covered small stocks, but not necessarily for large stocks (Guo and Mota, 2021). In comparison, when the optimism incentive is eliminated, competition monotonically improves aggregate information quality. This suggests that competition is only unambiguously good when optimism is sufficiently low, and a policy that increases both optimism and analyst competition could potentially lead to worse information in the financial market. Such a policy not only encourages analysts to be optimistic, but also potentially to *exaggerate* their optimism.[4]

The most recent round of policy changes enacts exactly this combination—increasing competition and optimism simultaneously—because reducing optimism and maintaining competition are currently framed as opposing objectives that must be traded off. My results challenge this framing, and show that due to interactions between the two incentives,

[4]Since the SEC rolled back the Settlement, observers have expressed concern that this could lead analysts to revert to issuing overly optimistic forecasts. Most notably, the SEC chair during the dot-com bubble, Arthur Levitt, warned that "The SEC May Make Wall Street Analysts Corrupt Again"; see `https://www.wsj.com/opinion/the-sec-may-make-wall-street-analysts-corrupt-again-79effdf4`. The concern became particularly salient after the SpaceX IPO, when analysts from all underwriters of the IPO issued bullish forecasts but not those from outside the syndicate, and the stock eventually underperformed those forecasts; see `https://www.ft.com/content/ce345155-d897-4f49-b7c8-13680e3b5434` and `https://www.profgmedia.com/p/the-analysts-are-compromised`. All links last accessed August 13, 2026.

information quality is maximized by accepting intermediate levels of competition. Policymakers should therefore consider more targeted regulations on analyst conflicts of interest.

Related literature This paper contributes to the growing literature on the industrial organization (IO) of financial markets, surveyed in Clark et al. (2021). It fits in the studies of intermediaries and agency problems (Bar-Isaac and Gavazza, 2015, Bhattacharya et al., 2025, Bolton et al., 2007, Chevalier and Ellison, 1997, 1999, Chu and Rysman, 2019, Egan, 2019, Egan et al., 2025, Robles-Garcia, 2020), adding to this literature by studying conflicts of interest in security analysts' information provision.

The model in this paper is not a standard empirical IO model, but a contest model in which players are rewarded for relative performance (Nalebuff and Stiglitz, 1983). I develop new methods to identify and estimate this model, exploiting variation in forecast strategy with competition. This is analogous to using variation in the number of bidders to test between common and private values in the auction literature (Athey and Haile, 2007, Compiani et al., 2020, Haile et al., 2003, Hendricks et al., 2003). This paper uses the same source of variation as the auction literature, but its model and estimation method depart from it.

This paper also adds to the literature on the effects of competition on information provision (e.g., Becker and Milbourn, 2011, Cagé, 2020, Gentzkow and Shapiro, 2008). One branch of this literature studies how competition affects experts' effort to acquire information (Dewatripont and Tirole, 1999, Pesendorfer and Wolinsky, 2003). This paper instead studies competition in the form of forecasting contests, holding the information environment fixed and focusing on experts' strategic choice of forecasts conditional on their signals; see Marinovic et al. (2013) and Bergemann and Ottaviani (2021) for surveys. In Section 3, I provide evidence that the forecasting-contest mechanism is likely the dominant one in my setting. I then generalize the model proposed by Ottaviani and Sørensen (2005, 2006) to account for optimism, a primary concern of policymakers in the security analyst market, and develop methods to structurally estimate this model.

Finally, this paper speaks to the broad literature on professional forecasters (e.g., Coibion and Gorodnichenko, 2012, Deb et al., 2018, Hong and Kubik, 2003). The behavioral finance literature uses analyst forecasts as measures of beliefs to study how beliefs respond to information shocks, abstracting from analysts' strategic interactions (Bordalo et al., 2019, Bouchaud et al., 2019).[5] I instead focus on these strategic interactions, studying how analysts choose forecasts in response to competition.[6] Several papers have studied

[5]Coibion and Gorodnichenko (2012) consider an extension where forecasters have a strategic incentive to report close to the average forecast.

[6]Relatedly, Camara (2015) estimates a structural model of forecast timing where analysts strategically

the impact of competition on analyst forecasts: Hong and Kacperczyk (2010) and Merkley et al. (2017) use brokerage mergers and closures, and Guo and Mota (2021) use MiFID II as exogenous shocks to analyst competition, but find mixed results on how reduced coverage affects forecast quality. By modeling the interaction between optimism and competition, this paper provides a mechanism where competition may have a positive or negative impact on information quality depending on the competitive environment.

I structurally estimate analysts' incentives in the model. A challenge to understanding these incentives is that analysts' payoffs are often unobservable, so existing studies proxy them with career outcomes (Clarke and Subramanian, 2006, Hong and Kubik, 2003).[7] This paper instead takes a revealed-preference approach, using analysts' forecast data to infer their incentives. The estimated model enables simulations of counterfactual policies that speak to current policy debates.

The remainder of the paper proceeds as follows. Section 2 introduces the background of security analysts, the policies on their conflict of interest, and the data. Section 3 provides reduced-form evidence to guide the setup of the model. Section 4 formally presents the model and characterizes the equilibrium, followed by Section 5, which establishes identification and details the estimation procedure and the results. Section 6 performs the counterfactual analyses and discusses their policy implications, and Section 7 concludes.

2 Industry Background and Data

2.1 The Sell-Side

Security analysts are typically employed by brokerage houses, commonly referred to as the sell-side. They are experts in the securities they cover and are connected to the securities' management teams, which allows them to obtain information. Using this information, they publish research reports on these securities, where they make earnings and stock-price forecasts, as well as buy, sell, or hold recommendations. The information they provide is crucial to investors.

This paper focuses on earnings forecasts, which offer an especially useful setting for studying analysts' strategic behavior in information provision for two reasons. First, firms disclose realized earnings following the end of each fiscal reporting period, so analysts'

choose when to release forecasts but always report their honest beliefs.

[7]Two exceptions are Groysberg et al. (2011) and Maber et al. (2021), which use proprietary compensation data from a single investment bank and find that analysts' compensation is significantly influenced by their rankings in "best-analyst" contests, though conditional on the contest outcomes, the effect of relative forecast accuracy is small.

forecasts can be compared with realized earnings to evaluate their quality ex post. Second, realized earnings are determined by firms' underlying profitability rather than forecasts. This feature sets them apart from other forecast targets, such as stock prices, which may be endogenous to analysts' forecasts and recommendations.

Brokerage houses generate revenue by executing trades for investors and providing investment-banking services to firms. Analysts' forecasts can increase a brokerage's revenue in two ways. Optimistic forecasts can stimulate trading volume and thereby increase brokerage commissions; they can also promote securities linked to the brokerage's investment-banking business (Jackson, 2005, Michaely and Womack, 1999). Accurate forecasts, on the other hand, build reputation for the brokerage (Jackson, 2005, Krigman et al., 2001), which may help attract future investment-banking business. Brokerages, therefore, could benefit from both the optimism and accuracy of their analysts' forecasts, and compensate their analysts accordingly.

Consistent with these incentives, analysts' payoffs depend on the optimism and accuracy of their forecasts. Analysts are typically not paid directly for individual research reports. Instead, their compensation consists of fixed wages, bonuses, and commissions, as well as intangible payoffs such as reputation, all of which are difficult to observe. Fixed wages largely depend on the prestige of the analyst's employer. High-prestige brokerages, including major investment banks such as JPMorgan, hire more analysts, take on larger deals, generate more revenue and therefore pay their analysts substantially more. Analysts with optimistic or accurate forecasts are more likely to move up this prestige hierarchy (Hong and Kubik, 2003). Bonuses and commissions, in turn, depend on the revenues that an analyst's forecasts help generate through trading and investment banking. Intangible payoffs such as reputation are even harder to observe. The unobservability of these various payoffs motivates my revealed-preference approach.

I model competition as a forecasting contest, where analysts are ranked and rewarded based on their forecast accuracy relative to their peers. This is an abstraction, but it is well grounded in the institutional setting. The financial industry often ranks experts based on their performance relative to their peers and disproportionately rewards top performers (Chevalier and Ellison, 1997). Security analysts, specifically, are ranked in a variety of "best analyst" awards, such as Institutional Investor's All-America Research Team and Thomson Reuters Analyst Awards, and winning these awards can lead to large increases in compensation and better job offers. Forecast accuracy enters these awards as either a direct or an indirect evaluation criterion to rank analysts. In addition, websites such as TipRanks.com publish rankings for analysts covering each security based on forecast accuracy, which are distributed by major financial data platforms such as Bloomberg.

2.2 Policies on Analysts' Conflict of Interest

Policies on analysts' conflict of interest typically operate through the two brokerage revenue sources: trading and investment banking. The SEC's regulations on favorable research after the dot-com bubble target the investment-banking channel, whereas MiFID II's unbundling rules target the trading channel. Interestingly, both sets of regulations were recently reversed, partly on the grounds of promoting competition, despite concerns about the resurgence of optimistic research due to analysts' conflict of interest. This policy uncertainty highlights the importance of understanding the interaction between competition and optimism, which motivates my counterfactual analyses.

After the dot-com bubble collapsed, the SEC introduced the regulation on favorable research in 2002, aiming to reduce analysts' optimism incentive. The regulation prohibits brokerages from tying analysts' compensation to specific investment-banking transactions and requires disclosure when compensation is linked to investment-banking revenues.[8] In 2003, the Settlement further imposed a $1.4 billion fine on ten of the largest US brokerages for fraudulent or exaggerated research and required them to separate their research and investment-banking departments.[9] However, in December 2025, the SEC reversed the Settlement by terminating its remaining undertakings, citing low analyst coverage for small stocks as one of the rationales, but industry commentators warned that this might revive analysts' optimism incentives.[10] MiFID II is a broad reform of the EU financial markets implemented in 2018. Among its most controversial policies are the unbundling rules, which require sell-side research to be charged separately from execution of financial instruments. Since 2024, these rules have been reversed both in the UK and in the EU, again with promoting competition as one of the main motivations.[11]

The Settlement was found to reduce optimistic stock recommendations and coverage at the affected brokerages, but its overall effects on earnings-forecast optimism and analyst competition remain ambiguous (Kadan et al., 2009, Merkley et al., 2017). MiFID II reduced analyst coverage for large stocks, while improving earnings-forecast quality and reducing trading-related optimism (Guo and Mota, 2021, Lourie et al., 2023). A mi-

[8] See `https://www.sec.gov/news/press/2002-63.htm`, accessed August 27, 2026.

[9] Two additional banks entered related settlements in 2004. See `https://www.sec.gov/news/press/2003-54.htm` and `https://www.sec.gov/news/press/2004-120.htm`, accessed August 3, 2026.

[10] See `https://www.sec.gov/newsroom/speeches-statements/uyeda-statement-global-research-analyst-settlement-120525` and `https://www.ft.com/content/ce345155-d897-4f49-b7c8-13680e3b5434`, accessed August 4, 2026.

[11] For the UK rules, see `https://www.fca.org.uk/publications/policy-statements/ps24-9-payment-for-optionality-investment-research`. For the EU, see Directive (EU) 2024/2811, `https://eur-lex.europa.eu/eli/dir/2024/2811/oj/eng`, applicable across member states from June 2026. All links were last accessed on August 4, 2026.

crofounded model of these policies requires modeling, at the very least, the trading and investment-banking businesses of brokerages and their connections to analysts' compensation, which are difficult to observe. Therefore, this paper instead estimates a "reduced-form" payoff for analysts to explicitly capture the optimism incentive and competition. The estimated model allows me to simulate the impact of policies that change both forces simultaneously and study their interaction.

2.3 Data

I obtain data on analysts' forecasts and employment history from the Institutional Brokers' Estimate System (IBES) Detail Earnings Estimate History. I focus on annual Earnings Per Share (EPS) forecasts of publicly listed US companies. For each forecast, IBES records both the forecast and the realized value, so I can measure forecast quality ex post, as well as when forecasts are issued and earnings are announced. IBES also records identifiers for securities, analysts, and brokerage houses, so I can trace each analyst's forecasts and employment history.

I match securities in IBES to the Center for Research in Security Prices (CRSP) database to obtain daily stock prices, stock returns, number of outstanding shares, and S&P 500 membership, and to Compustat to obtain financial information on the issuing firms such as book value and operating income. My sample covers forecasts announced between January 1, 1984 and December 31, 2016 for securities present in all three datasets. When an analyst revises her forecast, I keep the most recent forecast. I also follow the literature in excluding securities with forecast errors in the extreme tails to remove potential coding errors. Appendix B.1 presents the filtering details.

Finally, realized earnings evolve very differently across securities, varying in both volatility and trend. To make analysts' forecast problems comparable across securities, especially for structural estimation, I standardize them by assuming that analysts know the Markov process of the earnings. Then, in each security-year, analysts are effectively forecasting the standardized earnings shock based on their private signals. Appendix B.2 formalizes this standardization, after which analysts across all security-years have the same standard normal prior of the earnings. As a robustness check, Table A.2 shows that the main reduced-form results also hold using raw forecasts and realized earnings, without this standardization.

Table 1 reports the summary statistics. Panels A, B, and C present characteristics of analysts, brokerage houses, and securities, respectively. Now I highlight two facts.

Table 1: Summary Statistics

	Mean	Pctl(25)	Median	Pctl(75)
Panel A: Characteristics of Analysts				
Number of Securities Covered	6	2	5	8
Number of Industries Covered	2	1	1	2
High Prestige	0.513	0.000	1.000	1.000
Number of Lifetime Brokerage Houses	2	1	1	2
Number of Analyst-Year Observations		87,982		
Panel B: Characteristics of Brokerage Houses				
Number of Securities	67	5	20	71
Number of Industries	5	2	5	8
Number of Analysts	13	2	5	15
Number of Analysts per Security	1	1	1	1
Number of Brokerage-Year Observations		7,367		
Panel C: Characteristics of Securities				
$\log$(Market Cap)	13.549	12.292	13.469	14.730
$\log$(Book Value)	5.860	4.713	5.752	6.941
Profitability	0.173	0.086	0.194	0.294
Volatility ($\times 1{,}000$)	1.073	0.313	0.616	1.252
Average Monthly Return	0.013	-0.007	0.012	0.031
SP500	0.191	0.000	0.000	0.000
Share of High-Prestige Analysts	0.489	0.333	0.500	0.667
EPS Report Delay	42	28	40	53
Number of Security-Year Observations		41,336		

Note: Table reports summary statistics for the full sample. Panel A is measured at the analyst-year level, except that the number of lifetime brokerage houses is measured once per analyst over her observed career in IBES. Panel B is measured at the brokerage-year level. Panel C is measured at the security-year level. For a given security, *log(Market Cap)* is the natural logarithm of the security's fiscal-year-end market capitalization, calculated as the absolute CRSP month-end stock price multiplied by shares outstanding ($prc \times shrout$). *log(Book Value)* is the natural logarithm of the security's fiscal-year-end common book equity, calculated as Compustat total assets minus total liabilities and preferred stock ($at - lt - pstk$). *Profitability* is operating income after depreciation divided by common book equity ($oiadp/(at - lt - pstk)$). *Volatility* ($\times 1,000$) is the variance of daily CRSP stock returns over the 12 months ending in the fiscal-year-end month and is multiplied by 1,000 for readability. *Average Monthly Return* is the average CRSP monthly stock return over the same 12-month period. *SP500* is an indicator equal to one if the security is included in the S&P 500 on its fiscal-year-end date. *Share of High-Prestige Analysts* is the fraction of analysts covering the security in a fiscal year who are employed by high-prestige brokerage houses. A brokerage house is classified as high prestige if the number of analysts it employs is at or above the 90th percentile of brokerage houses in that year. *EPS Report Delay* is the number of days between the fiscal-year-end date and the earnings announcement date for that fiscal year. Industries are defined using the first two digits of the Thomson Reuters Business Classification (TRBC) permanent industry code; the sample spans 10 industries.

Analyst Expertise Each analyst specializes in a handful of securities concentrated in one or two industries. As Panel A of Table 1 shows, the median analyst covers only 5 securities and 1 industry in a year. This coverage is also highly persistent. Conditional on staying in the profession, 80% of analysts who cover a security in one year continue to cover it the following year. Once analysts stop covering a security, only 11% ever resume coverage. Such specialization and persistence suggest that analysts' coverage decisions are driven by their industry expertise and connections with the covered firms, rather than by temporary information shocks.

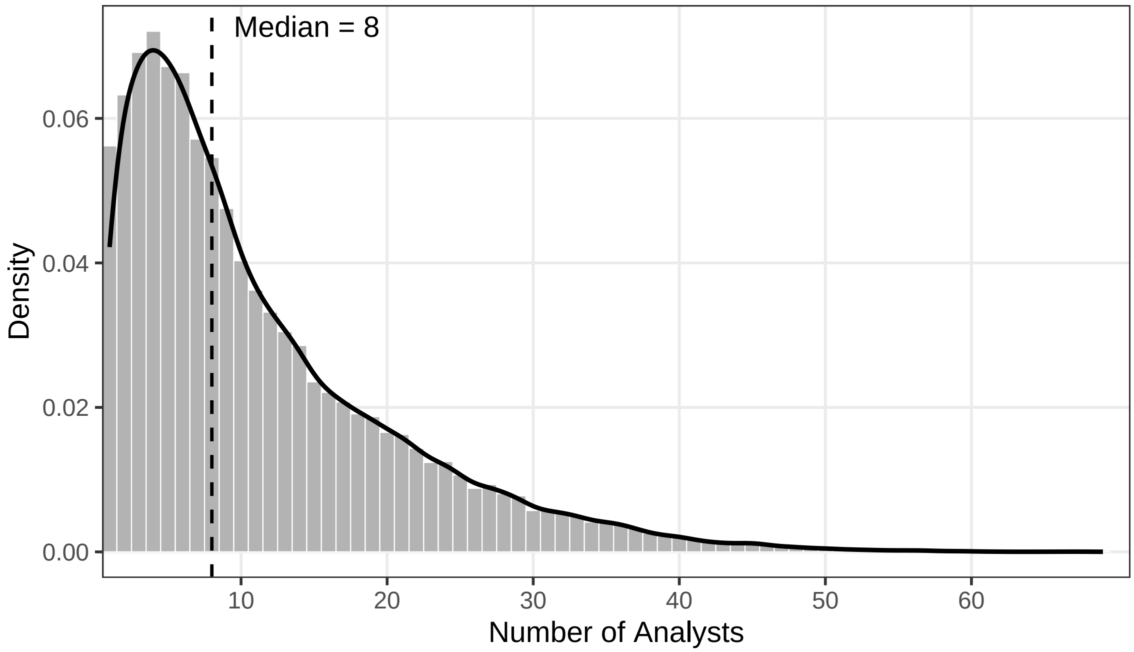
Competition I focus on analyst competition at the security level. Each security is considered a separate market, in which the covering analysts compete every year. Analysts' rivals come from other brokerage houses rather than their own, because each brokerage house typically assigns only one analyst to cover a given security-year (Table 1 Panel B). Figure 2 reports the distribution of competition, measured by the number of covering analysts, across security-years. It shows high levels of competition overall with a long right tail: the median security-year is covered by 8 analysts, and coverage reaches 69 analysts at the upper end of the distribution. Therefore, to stand out in a winner-takes-all contest in this industry, an analyst often needs to outperform a substantial number of rivals, which could create large distortionary incentives.

3 Reduced-Form Evidence

Now I document two data patterns that motivate the structural model. My analysis is organized around predictions about two extreme forms of relative-accuracy payoff: a reward for having the smallest forecast error (*winner-takes-all*) and a punishment for having the biggest forecast error (*loser-loses-all*). They both reward relative accuracy, but differ in payoff convexity with respect to rankings in the contest: a winner-takes-all payoff means that only the most accurate analyst gets a reward, so it is highly convex, whereas a loser-loses-all payoff means that only the least accurate analyst gets a punishment, so it is highly concave. The first prediction tests whether relative accuracy affects analysts' payoffs; the second uses the direction of the forecast distortions to show that analysts' behavior is more consistent with winner-takes-all. These extreme payoffs highlight the mechanism of analysts' strategic behavior; the structural model later allows for a more general payoff function.

Consider analysts forecasting the earnings of a given security-year. The analysts share a common prior over earnings, and each receives a private signal. After updating their

Figure 2: Distribution of Competition Across Security-Years



Note: Figure presents the distribution of competition across security-years in the full sample, where competition is measured by the number of analysts covering a security in a year. The bars form a density histogram with one-analyst-wide bins, and the solid line plots the corresponding kernel density estimate. The dashed vertical line marks the median.

beliefs, they make forecasts simultaneously. Each analyst's posterior mean falls between the prior mean and her private signal. Without strategic incentives, she would honestly report her posterior mean as her forecast. If she faces a winner-takes-all or loser-loses-all payoff, however, she would have an incentive to distort her forecast away from her posterior mean.

Prediction 1. Under either a winner-takes-all or loser-loses-all payoff, analysts make larger forecast errors when they face more rivals.

Prediction 2. Under a winner-takes-all payoff, analysts distort their forecasts towards their private signals. Under a loser-loses-all payoff, analysts distort their forecasts towards the common prior mean. In either case, the distortion increases with the number of rivals.

The intuition follows from the thought experiment in Figure 1. Suppose again that an analyst has a posterior mean of \$1, now arising from a prior mean of \$0 and a private signal of \$2. Under a winner-takes-all payoff, she has an incentive to deviate from her posterior mean of \$1 toward her private signal of \$2. If her forecast turns out to be correct, doing so increases the probability that she will be the *only* winner with the smallest forecast error, rather than tying with everyone else. By contrast, under a loser-loses-all payoff, she has an incentive to deviate from \$1 toward the prior mean of \$0, reducing the probability that she will be singled out and punished as the only loser with the largest forecast error. The more rivals she faces, the more she will need to deviate from \$1 in either direction to differentiate or hide herself (Prediction 2). And more deviation from her posterior mean leads to higher forecast error (Prediction 1).

Ottaviani and Sørensen (2005, 2006) establish these predictions analytically in a limiting case with the number of forecasters going to infinity, and numerically for finite numbers of forecasters. Later in Section 4, I extend their model to a more general payoff function and verify these predictions numerically using a newly developed algorithm for equilibrium computation. Here I take them as given to organize the reduced-form analysis.

3.1 Relative Accuracy Affects Analysts' Payoffs

Motivated by Prediction 1, I examine whether analysts' forecast errors increase with competition, as predicted when analysts face a winner-takes-all or a loser-loses-all payoff. For analyst i covering security j in year t, I estimate the panel regression

$$|X_{ijt} - Z_{jt}| = \alpha_1 \log(N_{jt}) + Controls_{ijt} + e_{1ijt}, \tag{1}$$

where X_{ijt} is analyst i's earnings forecast, Z_{jt} is the realized earnings, and $|X_{ijt} - Z_{jt}|$ is the forecast error. The variable N_{jt} is the number of analysts covering security j in year t. Prediction 1 implies that $\alpha_1 > 0$ when analysts face a winner-takes-all or a loser-loses-all payoff.

For the structural estimation, I retain security-years covered by at least two analysts, because relative accuracy is only defined when an analyst has a rival. I further focus on securities present in the data for at least 20 years, because the longer histories make it more plausible that analysts are informed about these securities and know their Markov processes, as assumed in Section 2.3. These securities are generally bigger, more profitable, less volatile, and covered by more analysts (Table A.1 and Figure A.2). I report the baseline reduced-form results using the same sample for consistency with the structural estimation. As shown below, the results are similar in the full sample and therefore are not driven by these sample restrictions.

Identification and instruments The main threat to identifying α_1 is that competition may be endogenous to a security's information environment. Securities that are harder to gather information on might attract fewer or more analysts, but analysts covering them might also have worse information and make larger errors. Year fixed effects absorb aggregate uncertainty and security fixed effects absorb permanent differences in complexity across securities. The more challenging case is if the information environment varies over time within a security in ways not captured by observables. I consider two natural mechanisms that could create this endogeneity.

The first mechanism is selection: analysts may select into covering certain securities. If the composition of analysts covering a security shifts with competition—for example, if highly covered securities also attract analysts who are worse at forecasting them—the estimated effect of competition would partly reflect who is covering rather than how many. This selection is the main mechanism documented by Guo and Mota (2021) in the context of MiFID II. To control for this, I include security-analyst pair fixed effects, so the effect of competition is identified from the *same* analyst covering the *same* security under different levels of competition. I also add time-varying controls for the security's information environment, the analyst's ability, and forecast timing.[12]

The second mechanism is that analysts switch in and out of covering a security in response to information shocks. To give a concrete example, an analyst who does not usually cover Amazon would have to pick it up one year because she stumbles upon some information, then drop coverage afterwards. This is unlikely to be the main source of variation for competition in this market, as analysts specialize in the firms they cover and rarely switch in and out of coverage (Table 1). Nevertheless, I address this concern with three instruments.

The first instrument is the lagged number of analysts. This instrument is relevant because coverage is persistent, and is exogenous if, conditional on the controls, lagged competition is uncorrelated with the current information shock. Because this assumption could fail if the confounding information shocks are serially correlated, I use two additional sources of variation. The second instrument uses brokerage-house mergers from Hong and Kacperczyk (2010). A brokerage typically assigns one analyst per security, so a merger removes duplicate coverage and reduces competition for securities covered by the merging parties. The mergers are unlikely to be driven by any security-level information shocks, because the merging parties are large financial institutions where equity research amounts to a small fraction of their business. The third instrument targets the switching concern directly. I freeze each analyst's coverage at the securities she covered in the first

[12]The security characteristics include market capitalization, book value, profitability, volatility, average monthly return, an indicator for inclusion in the S&P 500 index, and the share of analysts from high-prestige brokerages. The analyst characteristics include her experience, measured by the number of years she has covered the security, covered the industry, and stayed in the profession, along with an indicator for being employed at a high-prestige brokerage. The timing controls are the forecast's order and the earnings report delay. See the notes to Table 1 for details of variable construction.

Table 2: Effect of Competition on Analysts' Forecasts

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Forecast Error							
$\log(N_{jt})$	0.087	0.102	0.226	0.163	0.047	0.041	0.047
	(0.009)	(0.016)	(0.102)	(0.025)	(0.022)	(0.006)	(0.008)
Panel B: Forecast							
$\log(N_{jt}) \times Z_{jt}$	0.025	0.024	0.066	0.038	0.033	0.029	0.029
	(0.007)	(0.008)	(0.013)	(0.010)	(0.010)	(0.007)	(0.008)
Z_{jt}	0.810	0.812	0.692	0.777	0.791	0.794	0.793
	(0.020)	(0.021)	(0.040)	(0.028)	(0.028)	(0.020)	(0.021)
$\log(N_{jt})$	0.008	0.012	0.207	0.006	0.020	0.030	0.031
	(0.010)	(0.018)	(0.115)	(0.029)	(0.025)	(0.007)	(0.008)
Year FE	Y	Y	Y	Y	Y	Y	Y
Security FE						Y	Y
Security-analyst FE	Y	Y	Y	Y	Y		
Characteristics	Y	Y	Y	Y	Y		
Timing Controls	Y	Y	Y	Y	Y		
Lagged Competition IV		Y					Y
Merger IV			Y				
Early-coverage IV				Y			
Early-coverage-exit IV					Y		
Observations	195,892	195,892	195,892	143,092	142,863	209,320	209,320

Note: Table shows the effect of competition on analysts' forecasts. Panel A uses individual analysts' forecast errors as the dependent variable. Panel B uses individual analysts' forecasts as the dependent variable. Characteristics controls include security characteristics (market capitalization, book value, profitability, return volatility, average monthly returns, an indicator for inclusion in the S&P 500 index, and the share of covering analysts from high-prestige brokerage houses), analyst characteristics (experience with the security, with the industry, and with the profession, and prestige) and forecast-timing controls (forecast order and earnings-report delay). The lagged-competition IV is the one-year lag of the log number of analysts. The merger IV consists of 15 difference-in-differences-style indicators, one for each brokerage-house merger in the sample, each equal to one if and only if the security is covered by both merging brokerages, for fiscal years in or after the year of the merger. The early-coverage IV is, for each security-year, the log count of analysts who covered the security during their own first three years in the IBES sample and had entered the sample by that year; the early-coverage-exit IV additionally drops analysts after their final year in the sample. Both early-coverage columns exclude forecasts from analysts' first three years. First-stage strengths for all instruments are reported in Table A.3. The interaction term is instrumented with the interaction of the IV and the realized earnings. Standard errors are heteroskedasticity-consistent and clustered at the security-year level.

three years of her career and use the competition implied by these early-career choices, which cannot respond to any subsequent coverage changes. I construct two versions of this instrument that differ in whether analysts continue to be counted after they exit the profession. Table A.3 reports the Kleibergen–Paap F-statistic for each instrument as a measure of instrument strength.

Results Table 2 Panel A presents the results. Column (1) includes the security-analyst pair fixed effects, year fixed effects, and controls; Columns (2) through (5) add the lagged-competition, merger, early-coverage, and early-coverage-exit instruments, respectively. Forecast errors increase significantly with the number of analysts in every specification. Column (1) produces a coefficient of 0.087. In my preferred specification in Column (2), the coefficient on log number of analysts is 0.102. This implies that raising the number of analysts from 6 (the main-sample first quartile) to 12 (the main-sample median) increases forecast errors by 33%.[13]

The alternative instruments yield the same qualitative conclusion: the coefficients in Columns (3)–(5) are positive and range from 0.047 to 0.226. The merger estimate is the largest but also less precise, consistent with its relatively weak first stage.[14] These results are robust to both the sample restrictions and the standardization of forecasts and earnings. Table A.2 repeats the analysis in the full sample, which more than doubles the sample size and adds back smaller, more volatile, and less covered securities, using raw forecasts and realized earnings rather than their standardized counterparts. The coefficient on competition remains positive and statistically significant in every specification.

Connection to the structural model Columns (6) and (7) report simpler specifications with security and year fixed effects, without and with the lagged instrument. Their coefficients, 0.041 and 0.047, are smaller but preserve the qualitative relationship between competition and forecast errors. I report these specifications because they correspond most closely to the structural model, which assumes homogeneous analysts for tractability and uses similar regression coefficients, with security and year fixed effects concentrated out, as moments in estimation. The modestly larger coefficients in the richer specifications are consistent with analysts selecting into securities they understand better, which biases the coefficient on competition downward. Therefore, if anything, the simpler specifications are conservative about the importance of relative accuracy while preserving the key strategic incentive.

3.2 Winner-Takes-All or Loser-Loses-All

So far, I have shown that analysts' forecast errors increase when they face more competition, which is consistent with analysts responding strategically to rewards for relative accuracy. Now I distinguish between winner-takes-all and loser-loses-all payoffs.

As described in Prediction 2, under a winner-takes-all payoff, analysts differentiate themselves by moving their forecasts towards their private signals. Under a loser-loses-all payoff, in contrast, they move their forecasts away from their private signals and towards the common prior mean. In either case, the distortion intensifies as analysts face more

[13]The average forecast error with 6 analysts is 0.21. This percentage increase is computed as $33\% = [\log(12) - \log(6)] \times 0.102/0.21$.

[14]The merger specification has a Kleibergen–Paap F-statistic of 13.8, likely because the sample contains only 15 mergers.

rivals. The key to testing this prediction is to observe that all the information about the earnings shock is contained in analysts' private signals, so forecasts that put more weight on private signals should also respond more to changes in earnings. Concretely, I estimate the following panel regression:

$$X_{ijt} = \lambda_1 \log(N_{jt}) + \lambda_2 Z_{jt} + \lambda_3 \log(N_{jt}) \times Z_{jt} + Controls_{ijt} + e_{2ijt}. \tag{2}$$

This regression essentially tests whether analysts strategically underreact or overreact to their information about realized earnings, using variation in competition. A loser-loses-all payoff structure gives analysts a herding incentive that results in strategic underreaction to information, whereas a winner-takes-all payoff structure gives a contrarian incentive that results in strategic overreaction. Therefore, a positive coefficient on the interaction term ($\lambda_3 > 0$) implies that analysts respond more to information as they face more competition, which is more consistent with a winner-takes-all payoff.

Competition may be endogenous to the security's information environment for the same reasons as in Section 3.1, through selection or analysts switching in and out of coverage, so I adopt the same controls and instruments to address the endogeneity.

Results Panel B of Table 2 presents the results. The coefficient on the interaction between the realized earnings and the log number of analysts is positive and significant in every specification: when analysts face more rivals, their forecasts react more to the realized earnings, so they likely face a winner-takes-all payoff. In the preferred specification, Column (2), an increase in the number of analysts from 6 to 12 raises the coefficient on the realized earnings from 0.855 to 0.872 for the same analyst and the same security, implying that analysts react 2% ($= (0.872 - 0.855)/0.855$) more to shocks to realized earnings. The magnitude of the interaction is also stable across specifications, which suggests that heterogeneity across analysts plays a limited role within-security.

To summarize, I have presented two findings: when analysts face more competition, their forecasts have higher errors and react more to realized earnings. These findings are jointly consistent with a model where analysts' payoffs depend on relative accuracy, in a way that is closer to a winner-takes-all payoff, and analysts choose their forecasts strategically to maximize this payoff.

An alternative mechanism is that competition may affect analysts' information provision by influencing their effort to acquire information. If this were the dominant mechanism, however, we would expect the coefficient α_1 in equation (1) and λ_3 in equation (2) to have opposite signs. Suppose more competition induces analysts to exert more effort to acquire information, and thereby improves information quality. This should both reduce forecast errors and make forecasts react more to realized earnings. Because these two coefficients have the same sign across all specifications in Table 2, the forecasting contest is the more likely mechanism. Building on this evidence, I set up the structural model.

4 Model

I consider each security-year as a forecasting contest. For ease of exposition, the security subscript j and time subscript t are omitted in this section. A contest has $N \in \mathcal{N} \subset \mathbb{N}^{+}$ analysts, $i = 1, ..., N$. At the start of the contest, analysts share a common prior distribution of the earnings $Z \in \mathcal{Z} \subset \mathbb{R}$, but do not observe the realized earnings. They also observe the number of analysts in the contest N. Then, they draw private signals of the earnings, denoted by $S_i \in \mathcal{S} \subset \mathbb{R}$, from a continuous distribution, parameterized with τ. After that, they simultaneously make forecasts $X_i \in \mathcal{Z}$. Finally, the earnings Z are realized and analysts' payoffs are determined by a function $u : \mathcal{Z}^{N+1} \to \mathbb{R}$. Let $S = (S_1, ..., S_N)$, $X = (X_1, ..., X_N)$, $S_{-i} = S \backslash S_i$, and $X_{-i} = X \backslash X_i$. The payoff function u is defined as follows:

$$
\begin{aligned}
u(X_i, X_{-i}, Z) = &\underbrace{\sum_{k=1}^{N} \gamma_k(N) \mathbb{1}(\textit{analyst i has the k-th smallest forecast error})}_{\text{Relative Accuracy}} \\
&+ \underbrace{\gamma_c (X_i - Z)^2}_{\text{Absolute Accuracy}} + \underbrace{\gamma_o X_i}_{\text{Optimism}} .
\end{aligned}
$$

This payoff function has three components: relative accuracy, absolute accuracy, and optimism. It is parameterized by the vector γ, which determines the weight placed on each component.

Analyst i's forecast error is defined as the absolute difference between her forecast and the realized earnings $|X_i - Z|$. If an analyst has the k-th smallest forecast error, I say that she has rank k.[15] The payoff for relative accuracy is expressed with a series of weakly decreasing rewards in this rank: $\gamma_1(\cdot) \geq \gamma_2(\cdot) \geq \cdots \geq \gamma_N(\cdot)$. For example, $\gamma_1(\cdot)$ is the reward for being ranked first in the contest. These rewards may depend on N. Without loss of generality, I set the reward for the lowest-ranked analyst to be zero $\gamma_N(\cdot) \equiv 0$. This payoff for relative accuracy nests many commonly studied contests, such as:

[15] Analyst i has the k-th smallest forecast error when there are exactly k analysts j such that $|X_j - Z| \leq |X_i - Z|$.

- winner-takes-all: $\gamma_1(N) = \gamma_1 > 0$, $\gamma_k(N) = 0$ for $k \geq 2$;
- reward for top 10% analysts: $\gamma_k(N) = \gamma_1 > 0$ for $k \leq 10\% \times N$, otherwise $\gamma_k(N) = 0$;
- loser-loses-all: $\gamma_k(N) = \gamma_1 > 0$ for $k \leq N - 1$, $\gamma_N(N) = 0$.

I formulate this reward for relative accuracy flexibly and retain much of the flexibility when establishing the identification argument in Section 5.1. The specification used for structural estimation, however, is more parsimonious, because the endogeneity discussed in Section 3 limits the variation I can use to identify analysts' strategic incentives.

The next component, the payoff for absolute accuracy, is represented by a penalty for squared forecast error, γ_c, which I normalize to -1. By normalizing, I am implicitly assuming that analysts are always punished for forecast errors, which is plausible. If they were not, one might expect them to issue extremely high or low forecasts in response to the slightest incentive for optimism or pessimism, which is not observed in the data. The absolute accuracy component creates an *honest* scenario. If it is the only component in the payoff function, analysts will maximize payoffs by choosing the posterior means $\mathbb{E}(Z|S_i)$ as their forecasts, which I call the *honest* forecasts.

Finally, I model the payoff for optimism as a linear function of the forecast with coefficient γ_o. If $\gamma_o > 0$ and there is no payoff for relative accuracy, analysts will forecast higher than $\mathbb{E}(Z|S_i)$.

4.1 Equilibrium Existence

Analyst i's strategy is denoted by $\beta_i : \mathcal{S} \to \mathcal{Z}$. Given any set of rivals' strategies, β_{-i} , analyst i chooses forecast X_i to maximize the expected payoff conditional on receiving signal S_i:

$$U(X_i, S_i; \beta_{-i}(\cdot)) \equiv \int_Z \int_{S_{-i}} u(X_i, \beta_{-i}(S_{-i}), Z) h(S_{-i}, Z|S_i) dS_{-i} dZ, \tag{3}$$

where $h(S_{-i}, Z|S_i)$ represents the joint distribution of rivals' signals and the realized earnings conditional on analyst i's own signal.

Equilibrium β^* characterizes a pure-strategy Bayesian Nash equilibrium if and only if for all i and $S_i \in \mathcal{S}$,

$$\beta^*(S_i) = \arg\max_{X_i \in \mathcal{Z}} U(X_i, S_i; \beta^*_{-i}(\cdot)).$$

To guarantee the existence of an equilibrium, I assume that analysts' private signals are continuously distributed and correlated only through their correlation with the realized earnings. Formally,

Assumption 1 (Conditional Independence). $S_i \perp\!\!\!\perp S_j | Z$ for all $j \neq i$.

Then, I can express the conditional density of the realized earnings given i's private signal as $f(Z|S_i)$ and the conditional density of rival j's signal given the realized earnings as $g(S_j|Z)$. Equation (3) can be rewritten as

$$U(X_i, S_i; \beta_{-i}(\cdot)) \equiv \int_Z \int_{S_{-i}} u(X_i, \beta_{-i}(S_{-i}), Z) g(S_{-i}|Z) f(Z|S_i) dS_{-i} dZ. \tag{4}$$

Assumption 1 is sufficient for the existence of an equilibrium when the forecast space $\mathcal{Z}$ is finite, formalized in the following proposition.

Proposition 1. If Assumption 1 holds, analysts' private signals S_i are continuously distributed, and $\mathcal{Z}$ is finite, there exists a pure-strategy equilibrium.

Proof. See Appendix C.1. ■

Appendix C.2 extends this proposition to allow analysts to receive a common signal in addition to their private signals. An equilibrium exists as long as analysts' private signals are independent conditional on the realized earnings and the common signal, so the conditional independence assumption is not driving the main results of this paper.[16] Nonetheless, I maintain this assumption as it facilitates the identification of the structural model.

While the equilibrium existence is established for finite $\mathcal{Z}$, I will focus on a continuous state space in the rest of the paper, where the equilibrium is more tractable and faster to compute numerically, enabling estimation. Now I will characterize the equilibrium using first-order conditions and introduce the numerical procedure for equilibrium computation.

4.2 Equilibrium Characterization

The relative accuracy component of the objective function can be simplified as follows. First, consider analyst i and an arbitrary rival j. I say that j beats i when j's forecast is

[16] Intuitively, the common signal updates analysts' common prior distribution of the realized earnings. Once it is received, analysts will be playing the same contest with private signals, only with a new common prior, so all strategic interactions will remain the same.

more accurate than i's forecast, that is, $|X_j - Z| < |X_i - Z|$. This is equivalent to

$$\begin{aligned} X_i <& X_j < 2Z - X_i, \text{ if } X_i \leq Z \\ X_i >& X_j > 2Z - X_i, \text{ if } X_i > Z. \end{aligned}$$

I assume that analysts' forecast strategies are symmetric and strictly increasing in their private signals. This assumption allows me to infer analysts' private signals from their forecasts and guarantees the identification of the model (discussed further in Section 5.1).[17] Leveraging this, I write the probability of i beating one arbitrary rival j as

$$p(X_i, Z) \equiv 1 - \mathbb{P}(\textit{\textbf{j beats i}}|X_i, Z, S_i) = 1 - \left| \int_{\beta^{-1}(X_i)}^{\beta^{-1}(2Z - X_i)} g(S_j|Z) dS_j \right|.$$

Then, the probability of i having rank k is equal to the probability of i beating $N-k$ rivals and being beaten by $k-1$ rivals. The expected value of sending forecast X_i after observing signal S_i can be rewritten as

$$\begin{aligned} U(X_i, S_i; \beta_{-i}(\cdot)) = \int_{-\infty}^{\infty} \Big[\sum_{k=1}^{N} \gamma_k(N) \binom{N-1}{N-k} p(X_i, Z)^{N-k} (1 - p(X_i, Z))^{k-1} \\ - (X_i - Z)^2 + \gamma_o X_i \Big] f(Z|S_i) dZ. \end{aligned}$$

The first-order condition describing the symmetric and increasing equilibrium is given by

$$\begin{aligned} &\underbrace{\sum_{k=1}^{N} \gamma_k(N) \binom{N-1}{N-k} \int_{-\infty}^{\infty} \Xi(X_i, Z, N, k) f(Z|S_i) dZ}_{\text{Relative Accuracy}} \\ &\underbrace{-2[X_i - \mathbb{E}(Z|S_i)]}_{\text{Absolute Accuracy}} + \underbrace{\gamma_o}_{\text{Optimism}} = 0, \end{aligned} \tag{5}$$

to be solved for a strictly increasing $\beta^*(\cdot)$, where $X_i = \beta^*(S_i)$ and $\Xi(X_i, Z, N, k)$ is the

[17] While this is a strong assumption, it is verified numerically and also mirrors similar assumptions in the common-value auction literature to guarantee identification (Athey and Haile, 2007).

marginal probability of beating exactly a given set of $N-k$ rivals computed as

$$\begin{aligned}\Xi(X_i, Z, N, k) = &\big[\mathbb{1}(X_i \leq Z) - \mathbb{1}(X_i > Z)\big] \\ &\big[\frac{g(\beta^{*-1}(X_i)|Z)}{\beta^{*\prime}(\beta^{*-1}(X_i))} + \frac{g(\beta^{*-1}(2Z - X_i)|Z)}{\beta^{*\prime}(\beta^{*-1}(2Z - X_i))}\big] \\ &\big[\mathbb{1}(N > k)(N-k)p(X_i, Z)^{N-k-1}(1 - p(X_i, Z))^{k-1} - \\ &\mathbb{1}(k > 1)p(X_i, Z)^{N-k}(k-1)(1 - p(X_i, Z))^{k-2}\big]. \end{aligned} \tag{6}$$

The first-order condition has the same three components as the payoff function. The last two components are straightforward. The payoff for absolute accuracy creates an incentive for analysts to forecast close to their posteriors $\mathbb{E}(Z|S_i)$. Meanwhile, an optimism incentive implies that deviations above the posterior will bring higher expected payoff until the marginal loss from accuracy is higher than the marginal gain in optimism γ_o.

The relative accuracy component is the sum of the expected marginal reward for each rank. The marginal reward of being ranked k among N analysts is the reward $\gamma_k(N)$ multiplied by the marginal probability of receiving this rank, $\binom{N-1}{N-k} \int_{-\infty}^{\infty} \Xi(X_i, Z, N, k) f(Z|S_i) dZ$. Analysts choose forecasts to increase the probability of receiving better rewarded ranks. The first bracket of $\Xi(X_i, Z, N, k)$ in equation (6) shows that this can be further split into two terms, for when the realized earnings are higher than the forecast ($X_i \leq Z$), and when they are lower ($X_i > Z$). These two terms reflect the trade-off analysts face when they want to beat the rivals. A higher forecast improves the chance of beating the rivals when the realized earnings are higher than the forecast, but worsens the chance otherwise. The term $\big[\frac{g(\beta^{*-1}(X_i)|Z)}{\beta^{*\prime}(\beta^{*-1}(X_i))} + \frac{g(\beta^{*-1}(2Z-X_i)|Z)}{\beta^{*\prime}(\beta^{*-1}(2Z-X_i))}\big]$ represents the marginal gain or loss in the chance of beating one rival conditional on the realized earnings.

Prediction 2 introduces the intuition that when analysts are rewarded for the top, they will forecast close to their private signals, away from the common prior. Proposition 2 proves this formally in a winner-takes-all game with N players. I show that even without the payoff for optimism, reporting the posterior is not an equilibrium; rather, analysts have an incentive to deviate towards their signals.

Proposition 2. Suppose the realized earnings Z and the signals S are normally distributed and Assumption 1 holds. In a winner-takes-all game with N players, when analysts report their posterior means as forecasts, there exists a threshold $\bar{S}_N > 0$ such that for all $0 < |S_i| < \bar{S}_N$, an analyst's utility increases from a deviation in the direction of the signal.

Proof. See Appendix C.3. ■

4.3 Numerical Procedure for Equilibrium Computation

Now I introduce the numerical procedure used to solve the symmetric equilibrium. It is difficult to solve this game analytically, because the equilibrium is characterized by equation (5) for all signals in $\mathcal{S}$, which forms a complex system of differential equations involving derivatives of the equilibrium strategy. This system does not have a well-established solution to my knowledge. Therefore, I numerically approximate the equilibrium strategy using the projection method (Judd, 2020). It is fast to compute and flexible. Also, it allows me to focus on a symmetric equilibrium with a monotonic strategy, which is a difficult restriction to impose with an iterated best response algorithm.

I project the equilibrium strategy onto a monotonic polynomial parameterized with θ. First, I solve the derivative of the polynomial analytically. Second, I substitute the equilibrium strategy and its derivatives in the first-order conditions with those implied by the polynomial. Third, I find the θ^* that minimizes the sum of squared first-order conditions at a number of representative signals in $\mathcal{S}$ (also called "collocation points"). Finally, I verify whether the polynomial with θ^* is a good approximation of the equilibrium strategy by checking if the first-order conditions are indeed close to zero.

Under the parametric specifications of the structural estimation, which I will introduce in Section 5.2, I experiment with polynomials up to the 7th order and find a simple linear function, $X_i = \theta_1 + \theta_2 S_i$, to be an excellent approximation of the equilibrium strategy. This is also consistent with the analytical equilibrium strategies of the *honest* scenario and of the winner-takes-all game with an infinite number of players in Ottaviani and Sørensen (2006). Therefore, I adopt the linear approximation. Appendix D.1 presents the performance of polynomials up to the 7th order and additional details on the numerical procedure.

5 Identification, Estimation, and Results

The parameters of interest are the payoff function and the signal distribution. In Section 5.1, I establish sufficient conditions for nonparametric identification and show that these conditions are satisfied by many commonly studied contests. Then, I estimate the model using indirect inference.

The key to my identification strategy is to exploit variation in equilibrium forecasting behavior in response to competition, which identifies the payoff function. The intuition follows Predictions 1 and 2. In a winner-takes-all contest, for example, analysts make larger forecast errors when they face more rivals, and this response to competition is stronger when the reward for the most accurate analyst is larger. Thus, the change in forecast error with respect to competition is informative about the payoff function. As discussed in Section 3, I identify the structural model using only within-security, within-year variation in competition. This substantially limits the information available for estimation, so I impose additional structure on the model to obtain sufficiently precise estimates. Section 5.2 describes the estimation procedure, and Section 5.3 presents the results. Readers less interested in the theoretical details of the identification argument can proceed directly to Section 5.2.

5.1 Identification

I begin by showing that the signal distribution is nonparametrically identified, and then use the first-order conditions implied by the Bayesian Nash equilibrium in equation (5) to identify the payoff function.

5.1.1 Signal Distribution

The signal is only of information value in this model, so any strictly monotonic transformation of the signal does not change the information it contains (Athey and Haile, 2007). Consider the strictly increasing equilibrium forecast strategy under the true payoff $\beta^*(S_i; N)$ as this transformation. For a given number of analysts N, the set of signals can then be transformed to

$$\tilde{\mathcal{S}} \equiv \{\tilde{S} : \tilde{S} = \beta^*(S; N), S \in \mathcal{S}\}. \tag{7}$$

In other words, there is a bijection between the set of forecasts and the set of signals for a given N. Because we observe the forecasts, the signal distribution is nonparametrically identified.

This transformation also shows that both conditional distributions in the equilibrium first-order conditions are identified: 1) the density of the realized earnings conditional on the signal, $f(Z|S_i)$; 2) the density of the signal conditional on the realized earnings, $g(S_i|Z)$. For a given N, the density of the earnings conditional on the signal $f(Z|S_i)$ is equivalent to the density of the earnings conditional on the equilibrium forecast and the number of analysts, $\tilde{f}(Z|\beta^*(S_i; N), N)$. It is identified because Z, $\beta^*(S_i; N)$, and N are observed in the data. Similarly, $g(S_i|Z)$ is equivalent to $\tilde{g}(\beta^*(S_i; N)|Z, N)$, and thus identified as well.

5.1.2 Payoff Function

Given the identified signal distribution, I now establish a sufficient condition for nonparametric identification of the payoff function using the first-order conditions. Intuitively, the identification exploits how the equilibrium forecasting strategy responds to competition, and requires that this response have sufficient variation with every component of the payoff function.

Formally, I integrate the first-order condition in equation (5) over different regions of support for X, which has a continuous state space. For each $N \in \mathcal{N}$, choose $L(N) > 0$ intervals

$$I_l(N) = \big(\underline{x}_l(N), \overline{x}_l(N)\big), \qquad l = 1, \dots, L(N),$$

in the support of X. Following the transformation in equation (7), rewrite $f(Z|S_i)$ with the equilibrium forecast $X \equiv \beta^*(S_i; N)$ as $\tilde{f}(Z|X, N)$. Then, integrating equation (5) over any interval $I_l(N)$ gives a moment condition:

$$\begin{aligned}\gamma_o\left[\overline{x}_l(N) - \underline{x}_l(N)\right] + \sum_{k=1}^{N-1} \gamma_k(N) \binom{N-1}{N-k} \int_{\underline{x}_l(N)}^{\overline{x}_l(N)} \int_{-\infty}^{\infty} \Xi(X, Z, N, k) \tilde{f}(Z|X, N)\, dZ\, dX \\ = 2 \int_{\underline{x}_l(N)}^{\overline{x}_l(N)} \left[X - \mathbb{E}(Z|X, N)\right] dX. \end{aligned} \tag{8}$$

Equation (8) is linear in γ. Stacking these moment conditions across $l = 1, \dots, L(N)$ and $N \in \mathcal{N}$ gives a linear system $A\gamma = b$, where A denotes the Jacobian matrix.[18] I impose the following assumption on A as a sufficient condition for the global identification of γ.

[18] Specifically, let $\gamma(N) \equiv \big(\gamma_1(N), \dots, \gamma_{N-1}(N)\big)^T$, and let γ collect γ_o and $\gamma(N)$ for all N, ordered by increasing N:

$$\gamma \equiv \big(\gamma_o, \gamma(2)^T, \gamma(3)^T, \dots\big)^T.$$

Then, the Jacobian matrix A has $\sum_{N\in\mathcal{N}} L(N)$ rows and $\sum_{N\in\mathcal{N}}(N-1) + 1$ columns. Each interval $I_l(N)$ contributes one row to A. The row is zero except for the first element and the elements corresponding to $\gamma_k(N)$ for $k = 1, \dots, N-1$. The first element is

$$\overline{x}_l(N) - \underline{x}_l(N),$$

and the element corresponding to $\gamma_k(N)$ is

$$\binom{N-1}{N-k} \int_{\underline{x}_l(N)}^{\overline{x}_l(N)} \int_{-\infty}^{\infty} \Xi(X, Z, N, k) \tilde{f}(Z|X, N)\, dZ\, dX.$$

The corresponding row of the column vector b is

$$2 \int_{\underline{x}_l(N)}^{\overline{x}_l(N)} \left[X - \mathbb{E}(Z|X, N)\right] dX.$$

Assumption 2. There exists a collection of intervals

$$\{I_l(N) : l = 1, \ldots, L(N),\ N \in \mathcal{N}\}$$

such that the corresponding matrix A has full column rank.

Under Assumption 2, the system $A\gamma = b$ has a unique solution, so γ is identified. In addition, because the moment restrictions are linear in γ, the identification is global. This assumption requires sufficient variation in the first-order conditions across different numbers of analysts and equilibrium forecasts. In particular, the effects of different payoff components on the first-order condition must vary sufficiently across X and N to distinguish their parameters. As a more concrete example, I derive a more primitive sufficient condition for a special case where the reward for any rank of relative accuracy is homogeneous across different N's in Appendix C.4. I show that only mild conditions on the marginal probabilities of receiving different relative accuracy ranks are needed to guarantee identification in this case.

More importantly, Assumption 2 is testable because all components of A can be computed from the data. In equation (8), $\Xi(X, Z, N, k)$ is the marginal probability of receiving relative accuracy rank k with forecast X, conditioning on realized earnings Z; $\tilde{f}(Z|X, N)$ and $\mathbb{E}(Z|X, N)$ are the density and expectation of realized earnings, respectively, conditioning on equilibrium forecast X and the number of analysts N. Therefore, while it is beyond the scope of this paper to establish a general primitive condition that guarantees Assumption 2, one can test this assumption directly in the data for future applications.

5.2 Estimation Method

The parameters to be estimated are γ in the payoff function and τ in the signal distribution. I estimate them with indirect inference in four steps:

1. Run the auxiliary regressions in the data and use the regression coefficients as data moments (denoted $\hat{m}$).

2. Start with some initial parameter value $\{\tilde{\gamma}, \tilde{\tau}\}$. Compute the corresponding equilibrium strategy for all security-years in the data using the numerical procedure in Section 4.3.[19]

   An important assumption here is that if the model has multiple equilibria, all observations in the data are from the same equilibrium; otherwise, the model parameters

[19]To reduce the computation burden, the equilibrium strategy is computed for a grid of representative values of N and interpolated with cubic splines for the remaining values (Appendix D.1).

are not point identified.[20] In practice, the numerical procedure for equilibrium computation always gives the same symmetric monotone equilibrium.

3. Draw 100 simulations of analysts' private signals and use the equilibrium strategy to compute the forecasts. For each simulation, run the auxiliary regressions in the simulated data and use the average of the regression coefficients across simulations as the simulated moments, denoted by $m(\tilde{\gamma}, \tilde{\tau})$.

4. Match the simulated moments to the data moments to update $\{\tilde{\gamma}, \tilde{\tau}\}$. Given a weighting matrix W, repeat steps 2 and 3 until the weighted distance between the simulated moments and the data moments $(\hat{m} - m(\gamma, \tau))^T W(\hat{m} - m(\gamma, \tau))$ is minimized.

The identification strategy relies on the variation in the number of analysts, which could be endogenous due to unobserved confounders in the payoff or in the information environment as discussed in Section 3. Therefore, I concentrate out the differences at the security and year level to estimate a homogeneous contest. Analysts are assumed to be homogeneous otherwise, as the impact of cross-analyst heterogeneity within a security is limited. Following the specification in Table 2 Column (7), I first regress X_{ijt}, Z_{jt}, and N_{jt} on security and year fixed effects and take the residuals as the corresponding homogenized variables. Then, I instrument the number of analysts with the lagged number of analysts when running the auxiliary regressions. With a slight abuse of notation, X_{ijt}, Z_{jt}, and N_{jt} now denote the homogenized variables in the rest of the paper. Appendix D.2 presents the detailed homogenization procedure.

5.2.1 Parametric Specifications

Now I introduce the additional structure I impose to take the model to the data.

Signal Distribution The signal is assumed to be normally distributed around the realized earnings:

$$S_i = Z + \epsilon_i, \qquad \epsilon_i \sim N(0, 1/\tau),$$

where τ denotes the signal precision. Then, the posterior distribution of the realized earnings conditional on observing signal S_i is given by

$$Z|S_i \sim N(\frac{\tau S_i}{1+\tau}, \frac{1}{1+\tau}).$$

[20]This assumption is also used in dynamic oligopoly models and in first price auctions for which uniqueness has not been established (Athey and Haile, 2007, Bajari et al., 2007).

The posterior mean is the *honest* forecast, which puts more weight on the signal if the signal is more precise (i.e., τ is bigger).

Payoff Function The reduced-form evidence suggests that analysts' forecasts are more consistent with winner-takes-all than loser-loses-all payoffs (Table 2 Panel B). Motivated by this, I estimate a parsimonious baseline, where the relative accuracy component of the payoff function is winner-takes-all.

In addition, I estimate an alternative specification where the convexity of the relative accuracy reward is flexible, nesting winner-takes-all and loser-loses-all as limiting cases. This specification allows analysts ranked below the first to receive positive rewards and lets the data determine how concentrated the reward is at the top. The reward for rank k in a contest with N analysts is set to

$$\gamma_k(N) = \gamma_1 \left(1 - \frac{k-1}{N-1}\right)^{1/\gamma_{convex}},$$

where γ_1 is the reward for being the most accurate and γ_{convex} determines the convexity of the relative accuracy reward. Lower values of γ_{convex} imply a more convex reward. When γ_{convex} approaches zero, the reward approaches winner-takes-all; when $\gamma_{convex} = 1$, the reward is linear in rank; and when γ_{convex} approaches infinity, the reward approaches loser-loses-all. Appendix E formally proves properties of this specification and illustrates its shape.

5.2.2 Choice of Moments

Following the identification argument, I use the variation in forecast strategy in response to competition to identify the structural parameters. I estimate the following auxiliary regressions in the indirect inference procedure:

$$\mathbb{1}(X_{ijt} > Z_{jt}) = \tilde{\eta}_0 + \tilde{\eta}_1 \log(N_{jt}) + \varepsilon_{1ijt}, \tag{9}$$

$$|X_{ijt} - Z_{jt}| = \tilde{\alpha}_0 + \tilde{\alpha}_1 \log(N_{jt}) + \varepsilon_{2ijt}, \tag{10}$$

$$X_{ijt} = \tilde{\lambda}_0 + \tilde{\lambda}_1 \log(N_{jt}) + \tilde{\lambda}_2 Z_{jt} + \tilde{\lambda}_3 \log(N_{jt}) \times Z_{jt} + \varepsilon_{3ijt}. \tag{11}$$

These regressions largely mirror the reduced-form regressions in Section 3, but are estimated on the homogenized variables (Appendix D.2). Equation (10) corresponds to equation (1) and captures how the magnitude of analysts' forecast errors varies with competition. Equation (11) corresponds to equation (2) and captures how analysts' forecasts respond to information about realized earnings, and how this response varies with competition.

Compared with the reduced-form analysis, I include additional auxiliary regressions to capture the optimism incentive more directly. Equation (9) examines whether competition affects the probability of an optimistic forecast error. Specifically, $\mathbb{1}(X_{ijt} > Z_{jt})$ is an indicator equal to one when analyst i's forecast X_{ijt} is higher than realized earnings Z_{jt}. The coefficient on competition therefore captures how the probability of an optimistic forecast error varies with competition. I also estimate equation (10) separately for optimistic ($X_{ijt} > Z_{jt}$) and pessimistic ($X_{ijt} \leq Z_{jt}$) forecast errors, which allows me to assess whether the effect of competition on the magnitude of forecast errors differs depending on whether forecasts are above or below realized earnings.

I use the coefficients $\tilde{\eta}, \tilde{\alpha}, \tilde{\lambda}$, and the forecast residual variance from equation (11) as moments. The weighting matrix W is a block-diagonal matrix. Each block corresponding to an auxiliary regression is given by the inverse of the sample variance-covariance matrix with heteroskedasticity-consistent standard errors clustered at the security-year level. The forecast residual variance is assigned a weight equal to the average weight of the remaining moments.

5.3 Results

Table 3 presents the estimation results. Column (1) reports the winner-takes-all baseline. I find a large and statistically significant reward for being the most accurate analyst. The magnitude of this reward is equivalent to the absolute accuracy penalty of a forecast that is 3.4 $(= \sqrt{11.719})$ standard deviations away from the realized earnings. I also find a significant incentive for optimism. The estimated optimism coefficient implies that, absent the relative accuracy reward, an analyst receives a higher payoff from issuing an optimistic forecast up to 0.955 standard deviations above the realized earnings, than from reporting the realized earnings itself.

Column (2) relaxes the winner-takes-all restriction and allows the data to determine the convexity of the relative accuracy reward. The estimated shape parameter is 0.020, which is close to zero. This suggests that the relative accuracy reward is close to winner-takes-all. Panel B further translates this estimate into the implied reward schedule for a security-year with eight analysts, the median in the full sample (Figure 2). The second-ranked analyst receives 0.04% of the reward for being the most accurate, the third-ranked analyst receives less than 0.01%, and the most accurate analyst accounts for 99.96% of the total relative accuracy reward. Therefore, the estimated reward is still very concentrated

Table 3: Estimates and Implied Relative-Accuracy Rewards

	(1) Winner-takes-all	(2) Flexible Convexity
Panel A: Structural parameters		
Reward for the Most Accurate, γ_1	11.719	10.823
	(3.092)	(4.000)
Optimism, γ_o	0.955	1.134
	(0.101)	(0.217)
Signal Precision, τ	5.297	5.178
	(0.169)	(0.164)
Reward Shape, γ_{convex}	—	0.020
		(0.005)
Panel B: Implied reward schedule at $N = 8$		
Rank-2 Reward / Rank-1 Reward (%)	0.00	0.04
Rank-3 Reward / Rank-1 Reward (%)	0.00	< 0.01
Rank-1 Share of Total Rank Rewards (%)	100.00	99.96

Note: Table presents estimates of analysts' payoff function and signal precision under alternative specifications of the relative accuracy reward. The winner-takes-all specification rewards only the most accurate analyst. The flexible specification sets the reward for rank k to $\gamma_k(N) = \gamma_1[1 - (k-1)/(N-1)]^{1/\gamma_{convex}}$. The coefficient on absolute accuracy is normalized to -1 in both specifications. Panel B reports the implied reward schedule at $N = 8$, the median number of analysts per security-year in the full sample. Asymptotic simulated method of moments standard errors, based on market-clustered moment covariances and adjusted for simulation error, are reported in parentheses.

on the most accurate even when the convexity is estimated. The remaining estimates are also consistent across the two specifications. Therefore, I proceed to use the winner-takes-all baseline to evaluate the fit of the model and conduct the counterfactual analyses.

Table 4 presents the fit of the auxiliary regression coefficients on the key regressors. The model fits the data reasonably well. The structural estimates reproduce the two reduced-form patterns: analysts' forecast errors increase with competition, and their forecasts respond more to the realized earnings when they face more rivals. Consistent with the data, the simulated probability of reporting an optimistic forecast also increases with competition. One discrepancy is that the effect of competition on pessimistic forecast errors has the opposite sign in the data and the simulation, although this moment is close to zero in both.

6 Counterfactuals

Using the estimated model, I simulate counterfactuals to study how optimism and competition jointly shape information quality. Understanding their interaction is crucial to evaluating policies on analysts' conflict of interest, which tend to affect both at the same time.

Table 4: Moment Fit

	Data	Model
Panel A: Optimism regression		
Log number of analysts, $\log(N_{jt})$	0.085	0.086
Panel B: Forecast-error regressions		
All forecasts	0.047	0.034
Optimistic forecasts, $X_{ijt} > Z_{jt}$	0.068	0.050
Pessimistic forecasts, $X_{ijt} \leq Z_{jt}$	0.026	-0.021
Panel C: Forecast regression		
Log number of analysts, $\log(N_{jt})$	0.042	0.095
Realized earnings, Z_{jt}	0.833	0.832
$Z_{jt} \times \log(N_{jt})$	0.015	0.012

Note: Table compares the auxiliary regression coefficients on the key regressors in the data and in the simulated sample under the winner-takes-all baseline. Panel A reports the coefficient on $\log(N_{jt})$ from the optimism regression in equation (9). Panel B reports coefficients on $\log(N_{jt})$ from the forecast-error regression in equation (10) for all forecasts and separately for optimistic and pessimistic forecasts. Panel C reports coefficients from the forecast regression in equation (11).

I start with an overview of the equilibrium strategies under different payoffs in Section 6.1. Section 6.2 then eliminates the relative accuracy and optimism rewards separately or simultaneously, holding the observed competitive environment fixed, to quantify their marginal effects on information quality. Finally, Section 6.3 varies the level of competition to quantify its impact on information quality.

For a given payoff function, I simulate analysts' forecasts as follows. For each contest, I draw a random number from the standard normal distribution as the realized earnings. Then, for each analyst, I draw a random noise, normally distributed with mean zero and variance $1/\tau$, and add it to the realized earnings to create a signal. Finally, I solve the equilibrium strategy for each contest to simulate forecasts from the signals.

I evaluate the quality of these forecasts at face value, assuming that investors do not know enough about analysts' incentives to back out their private signals from the forecasts. This assumption may be strong for institutional investors, who might understand sell-side incentives and discount analysts' forecasts accordingly. But retail investors might not share the same understanding and hence may be less able to do so. This inability to fully discount analysts' forecasts is reflected in the dot-com bubble, when investors suffered substantial losses from biased analyst research, which eventually led to the Settlement. Prior evidence also shows that investors do not fully undo the optimism in analysts' forecasts (Bordalo et al., 2024). Hence, I proceed under the face-value assumption.

To evaluate information quality, I consider both individual and consensus forecasts. The consensus forecast is defined as the mean individual forecast for a given security-year, which captures aggregate information in the market.[21] It is publicly available through major financial news outlets and frequently cited by firms and investors. Therefore, policymakers would likely seek to optimize the quality of the consensus, which will be the focus of my discussions. For reference, I report individual forecast quality to demonstrate the underlying mechanism. Effects on individual and consensus quality often point in the same direction, with a few exceptions. The main point of divergence, intuitively, is that the consensus benefits from analysts incorporating more private information into their forecasts even if it is noisy, because the noise may "cancel out" to some extent. Appendix C.5 formalizes and proves this as a lemma. I measure the accuracy and noisiness of information provided by any individual analyst using the average forecast error and the dispersion of individual forecasts within each security-year, respectively. For aggregate information, I use the average forecast error and the variance of consensus forecasts across security-years.

6.1 Equilibrium Strategy

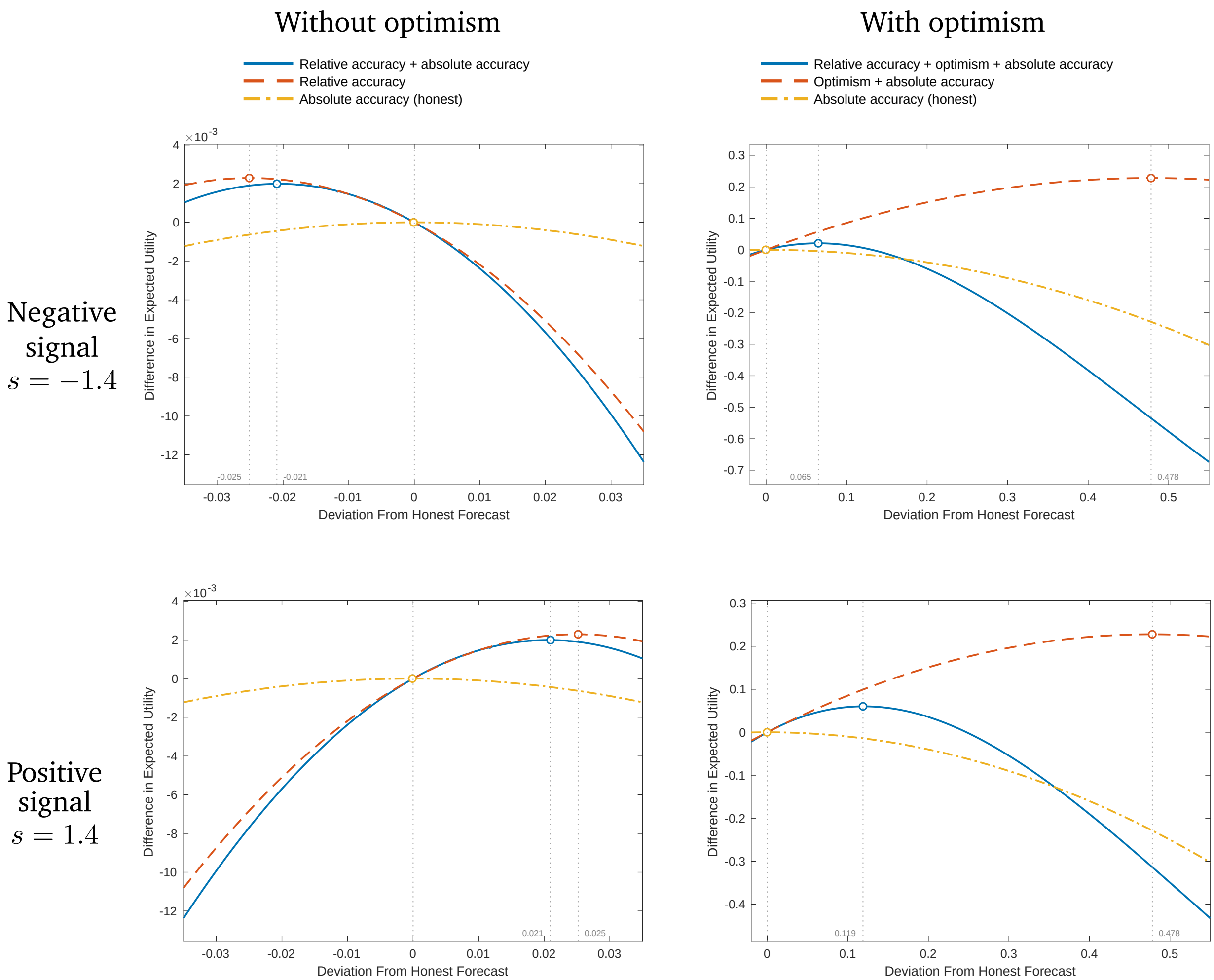
I start with an example to illustrate analysts' equilibrium forecasting strategies. Figure 3 plots the gains and losses in expected utility when an analyst deviates from her honest forecast under different payoffs—either the estimated payoff, or the estimated payoff with some components shut down. The analyst competes with 9 rivals, who play the corresponding equilibrium strategy under each payoff. She receives either a negative signal of -1.4 (top panels) or a positive signal of 1.4 (bottom panels), which is at the 10th or 90th percentile of the signal distribution and implies an honest forecast of -1.2 or 1.2, respectively.

The left panels present the payoffs with optimism eliminated. When analysts are rewarded for relative accuracy but not optimism, they deviate in the direction of their signals to differentiate themselves: those with positive signals distort their forecasts upwards, and those with negative signals distort their forecasts downwards. The magnitude of the distortion is slightly bigger when relative accuracy is the only component in the payoff, but still small, only about 2% ($= 0.025/1.2$).

The right panels add back the estimated optimism incentive. This optimism would have resulted in huge distortions upwards regardless of analysts' signals, if the reward for relative accuracy were absent. This relative accuracy reward disciplines this behavior so

[21] The consensus forecast is also sometimes computed as the median of individual forecasts, which yields similar counterfactual results.

Figure 3: Expected Utility Under Alternative Incentives and Private Signals



Note: Figure plots an analyst's expected utility at the estimated parameters with 9 rivals playing the corresponding equilibrium strategy. The analyst receives a negative signal of $s = -1.4$ in the top panels and a positive signal of $s = 1.4$ in the bottom panels, which are at the 10th and 90th percentiles of the signal distribution, respectively. The horizontal axis shows the deviation from the honest forecast, $X_i - \mathbb{E}(Z \mid S_i = s)$, while the vertical axis shows the difference in expected utility between reporting a given forecast and the honest forecast. In the left panels, the blue solid line includes relative and absolute accuracy in analysts' payoff function, the red dashed line includes only relative accuracy, and the yellow dash-dot line includes only absolute accuracy. In the right panels, the blue solid line includes relative accuracy, optimism, and absolute accuracy, the red dashed line includes optimism and absolute accuracy, and the yellow dash-dot line includes only absolute accuracy. The circles indicate the equilibrium forecast under each payoff specification and their deviations are marked in gray above the x-axis.

Table 5: Marginal Effects of Payoff Components on Information Quality

	Individual		**Consensus**	
Payoff Change	Forecast Error	Forecast Dispersion	Forecast Error	Forecast Variance
Add optimism	↑ 8.36%	↑ 3.32%	↑ 39.40%	↑ 3.41%
Add relative accuracy	↓ 34.01%	↑ 6.59%	↓ 60.84%	↑ 6.68%
Both incentives	↑ 8.43%	↑ 6.59%	↑ 30.89%	↑ 6.68%

Note: Table shows the marginal effects of adding optimism, relative accuracy, or both components to analysts' payoff function. The first row adds optimism to a payoff function that rewards relative and absolute accuracy. The second row adds relative accuracy to a payoff function that rewards optimism and absolute accuracy. The third row adds both optimism and relative accuracy to a payoff function that rewards only absolute accuracy. Each cell reports the percentage difference in information quality relative to the payoff structure without the added component(s). The arrows indicate whether the measure increases or decreases. For a given contest, individual forecast error is $\sum_i |X_i - Z|/N$, individual forecast dispersion is $\sum_i (X_i - \bar{X})^2/(N-1)$, and consensus forecast error is $|X^* - Z|$, where X^* denotes the consensus forecast. The table reports the mean of these measures across simulated contests. The variance of the consensus forecast is the variance of X^* across simulated contests.

that analysts report less optimistic forecasts. Most of them still distort upwards, just by a smaller amount, because the optimism incentive is strong. The plotted analyst with the positive signal of 1.4 would distort her forecast up by 40% ($= 0.48/1.2$) in the absence of the relative accuracy reward, but by only 10% ($= 0.12/1.2$) with the reward.

A comparison of the top right and bottom right panels shows that the interaction between the forecasting contest and the optimism incentive creates an asymmetric effect: analysts with positive signals distort more than those with negative signals. Analysts with positive signals are pushed by the rewards for relative accuracy and optimism to distort in the same direction. They know that their rivals are optimistic, so they need to be *even more bullish* to stand out. For analysts with negative signals, on the other hand, relative accuracy and optimism lead them to distort in opposite directions, so the two forces counteract each other. This asymmetry magnifies noise in analysts' forecasts, as bullish analysts disproportionately exaggerate their signals. I will elaborate further as I discuss the marginal effect of each payoff component below.

6.2 Incentive Decomposition

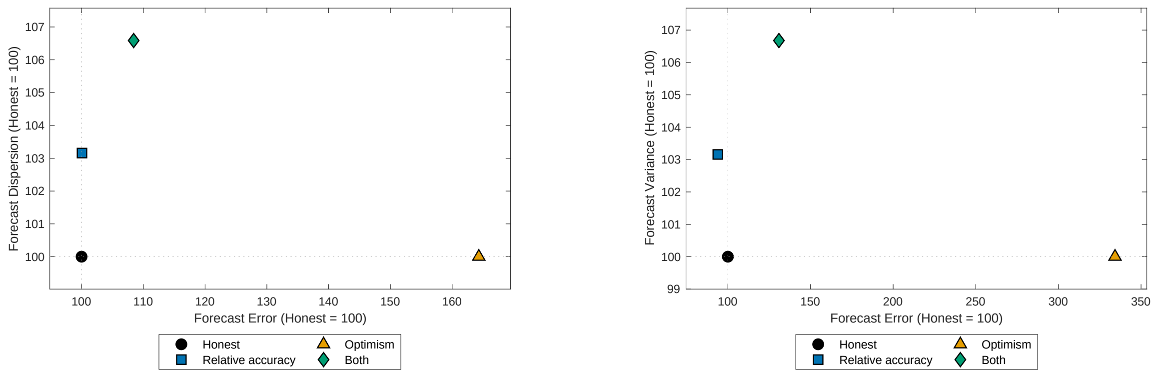
How do these equilibrium strategies translate into information quality? To quantify this, for any given payoff function, I simulate every security-year in the data 20 times and average the information quality measures across simulations and observations. I compute the marginal effect of each payoff component as the percentage change in information quality from adding that component. Table 5 presents the marginal effect of optimism, relative accuracy, and both jointly.

First, optimism increases not only forecast error but also forecast noise. Adding the optimism incentive increases individual and consensus forecast errors by 8.36% and 39.40%,

Figure 4: Trade-off Between Forecast Accuracy and Noise

(a) Individual Forecasts

(b) Consensus Forecast



Note: Figure shows the trade-off between forecast accuracy and noise under four payoff structures. Panel (a) plots individual forecast error against individual forecast dispersion, while panel (b) plots consensus forecast error against the variance of the consensus forecast. Each measure is normalized by its value when analysts are rewarded only for absolute accuracy, so the honest baseline equals 100 on both axes. "Relative accuracy" then adds the reward for relative accuracy; "Optimism" adds the reward for optimism; and "Both" adds the rewards for relative accuracy and optimism. The horizontal and vertical dotted lines mark the honest baseline. For a given contest, individual forecast error is $\sum_i |X_i - Z|/N$, individual forecast dispersion is $\sum_i (X_i - \bar{X})^2/(N-1)$, and consensus forecast error is $|X^* - Z|$, where X^* denotes the consensus forecast. The figure reports the mean of these measures across simulated contests. The variance of the consensus forecast is the variance of X^* across simulated contests. A point farther to the left represents lower forecast error, while a point lower in the figure represents less noise.

respectively, because it leads analysts to report overly optimistic forecasts. It also raises the noise measures by about 3% because it interacts with the relative accuracy reward to create the asymmetric effect. As a result, analysts with positive signals exaggerate their signals disproportionately more, amplifying the noise in the market.

Second, the forecasting contest disciplines optimism, but at the cost of increasing noise. The reward for relative accuracy reduces individual and consensus forecast errors—which are large in the first place because of optimism—by 34.01% and 60.84% respectively. This reduction in forecast errors shows that the disciplinary effect dominates, so from an information accuracy perspective, contests are net positive in the current market. Nevertheless, the distortionary effect is still there. The reward for relative accuracy still gives analysts an incentive to differentiate themselves and its interaction with optimism amplifies noise in the market, increasing it by about 7%. I have been highlighting this same interaction but with different baselines. Relative accuracy alone amplifies forecast noise as analysts differentiate themselves by moderately exaggerating their signals, and when it is combined with optimism, this exaggeration is coupled with the asymmetric effect, so the noise is even greater.

The rewards for relative accuracy and optimism jointly increase both forecast error and noise. Figure 4 visualizes the trade-off between these two dimensions of information quality.[22] In isolation, the optimism reward has a greater impact on forecast error, whereas the relative accuracy reward has a greater impact on noise (holding the reward for absolute accuracy fixed). The interaction between the two leads to more noise compared to either incentive alone. However, if policy tools cannot eliminate the optimism incentive—whether because enforcement is hard or policy uncertainty is high, as industry observers warned after the recent policy reversals—then the relative accuracy reward at least serves as a "market-based" mechanism to discipline optimism, albeit at the cost of greater noise in the financial market.

6.3 Effect of Competition

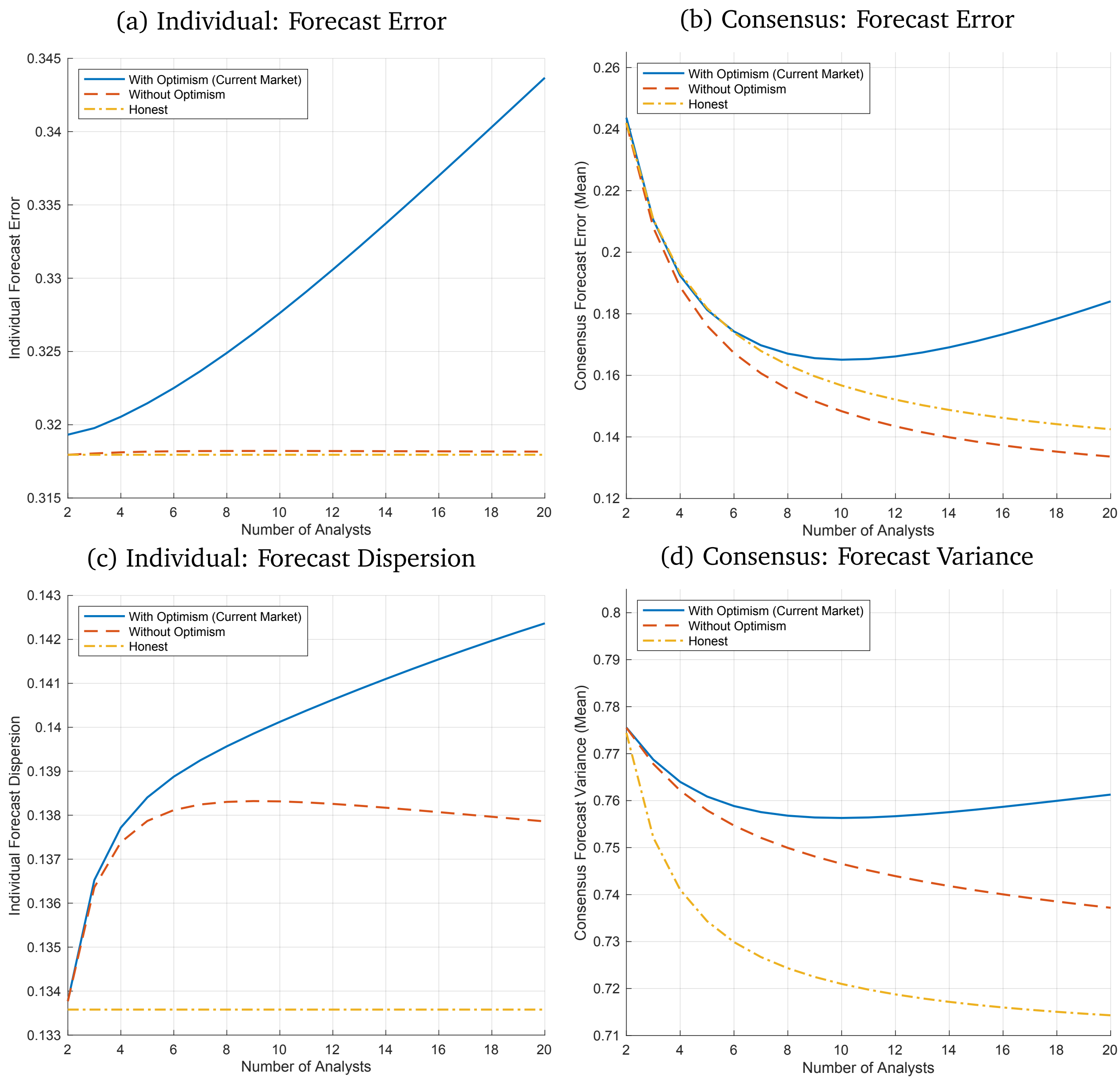
The analysis so far has held the competitive environment fixed. Now I turn to the effect of competition on information quality. Figure 5 plots how information quality changes with competition under different payoffs. For each payoff and competition level, I solve for the equilibrium forecasting strategy numerically using the projection method, and then use it to analytically derive the information-quality measures. Appendix C.6 presents the detailed expressions and their derivations.

Panels (a) and (b) of Figure 5 show how information accuracy as measured by forecast errors changes with competition. When analysts are only rewarded for absolute accuracy, they report their honest beliefs, so the individual forecast error is constant with respect to competition (yellow dash-dot line). Consequently, the consensus forecast error decreases with competition because there is more aggregate information, or put differently, because of the Law of Large Numbers. When analysts are rewarded for relative accuracy in addition (red dashed line), the individual forecast error becomes only slightly higher than that of the honest forecasts and mildly increasing with competition up to 9 analysts, so the consensus forecast error still decreases with competition.

Finally, when analysts are further rewarded for optimism and face the complete estimated payoff, the individual forecast error increases with competition (blue solid line), replicating one of the key variations that I use for identification. The contrast between this and the line without optimism shows that the large effect of competition on forecast error in Section 3 is predominantly due to the interaction between optimism and relative accuracy. This interaction leads analysts, especially those with positive signals, to make

[22] One interesting observation here is that the consensus forecast error is lower with relative accuracy than under the honest payoff. As discussed earlier, this is because the consensus forecast error (with more than one analyst) is minimized when each analyst puts more weight on her private signal than she does in her honest forecast. In other words, the consensus forecast benefits from analysts revealing slightly more of their own signals, which is precisely what analysts do when they face the relative accuracy reward. As a result, the consensus forecast error falls.

Figure 5: The Effect of Competition



Note: Figure shows measures of information quality under different payoff specifications and competition levels. Panel (a) plots mean forecast error of individual forecasts; panel (b) mean forecast error of the consensus forecast; panel (c) mean forecast dispersion of individual forecasts; panel (d) variance of the consensus forecast. The blue solid line represents the current market with relative accuracy, optimism, and absolute accuracy in analysts' payoff function. The red dashed line includes relative and absolute accuracy, while the yellow dash-dot line includes only absolute accuracy.

much larger positive distortions, so the distortionary effect becomes more pronounced. Consequently, the consensus forecast error is U-shaped in the level of competition: it first decreases with competition thanks to more aggregate information, but as competition intensifies, it starts to increase as the distortionary effect becomes more pronounced and overtakes the benefit of more aggregate information.

Panels (c) and (d) of Figure 5 show similar patterns with the noise measures: under the estimated payoff, the consensus forecast variance is U-shaped in the level of competition. The consensus quality in terms of both accuracy and noise is optimized at 10 analysts per security-year, which is at the 59th percentile of the full-sample distribution. This implies that over 40% of the security-years lie beyond this threshold, where additional analyst competition could reduce consensus quality.

These results challenge two common beliefs about information provided by security analysts. First, analysts' forecast dispersion is often used as a measure of uncertainty in the market, but my results suggest that it can also be driven by analysts' strategic incentives. The level of competition should be taken into account when using forecast dispersion as an uncertainty measure. Second, it is commonly believed that more competition is always better for information quality, but I show that when analysts strategically compete with each other, this might no longer be the case. The interaction between optimism and competition makes the consensus quality inverted-U-shaped with respect to the level of competition: moderate competition disciplines optimism, but too much competition produces information that is less accurate and noisier.

6.4 Policy Implications

What do these results imply about policies on analysts' conflict of interest? Unlike in many other markets, where competition can be regulated through antitrust policies, direct regulation of analyst competition is likely infeasible, because analysts are not firms but experts employed by brokerages. Mergers and acquisitions between brokerages shift analyst competition across many stocks at the same time, each with a different preexisting competitive environment, so we cannot prescribe a universal optimal level of competition.

However, this does not mean that competition is not a core concern in policymaking. On the one hand, the concern about low analyst coverage of small stocks, which motivated the recent reversals of the Settlement and the MiFID II unbundling rule, is well founded. I show that, regardless of whether analysts face the optimism incentive, increasing competition for less-covered stocks consistently improves consensus quality.

On the other hand, these reversals would likely increase both optimism and competition. The conventional argument is that competition always improves information, so even if the additional optimism is undesirable, the net effect of these reversals might still be positive. However, my simulations suggest that for highly covered stocks, bullish analysts may disproportionately exaggerate their signals, producing forecasts that are substantially less accurate and noisier. It is precisely the interaction between competition and optimism that could ultimately reduce information quality. In the current equity market, where a large and growing share of aggregate market capitalization is concentrated in a small number of high-profile, heavily covered stocks, policymakers should be particularly cautious about such policies that strengthen both forces and consider more targeted regulations to expand coverage of smaller stocks.

7 Conclusion

This paper studies how competition affects information quality provided by security analysts. I build and estimate a forecasting contest model, in which analysts strategically choose their forecasts in response to incentives for absolute accuracy, relative accuracy, and optimism. Policymakers often perceive a trade-off between mitigating optimism and preserving competition as they regulate analysts' conflict of interest. My model incorporates both forces in a unified framework, and reveals that their interaction is more nuanced than a simple trade-off. A challenge in understanding analysts' incentives is that their payoffs are often unobservable. To address this, this paper takes a revealed-preference approach. I develop a new identification strategy and structurally estimate analysts' payoffs from their forecasts to infer their incentives.

I find that analysts receive significant rewards for being the most accurate and for optimism. The estimated payoff implies that, in equilibrium, nearly all analysts distort their forecasts upwards due to the large optimism incentive, while the reward for the most accurate gives analysts an incentive to differentiate themselves from their rivals. Their interaction leads to asymmetric distortions: analysts with positive signals distort upwards disproportionately more than those with negative signals.

As a result, the optimism incentive not only increases analysts' forecast errors as they become more biased, but also amplifies noise as bullish analysts disproportionately exaggerate their signals. The aggregate information quality, as reflected in consensus forecasts, becomes inverted-U-shaped with respect to competition. At low levels of competition, additional competition increases the amount of aggregate information in the market and hence improves information quality. But as competition intensifies, analysts distort their forecasts more in order to differentiate, so information quality deteriorates. The reward for the most accurate serves a dual role: it makes analysts distort their forecasts to stand out, but also disciplines them against issuing overly optimistic forecasts. On net, the disciplinary effect dominates: rewarding the most accurate analyst reduces forecast errors, serving as a "market-based" mechanism to curb the optimism incentive, but this is at the cost of making the market noisier.

These results have important policy implications. In the security analyst market, competition is often viewed as unambiguously beneficial for information quality, but my findings show that its effects depend on both the magnitude of the optimism incentive and the competitive environment. When a stock is already heavily covered and analysts face strong incentives to issue optimistic forecasts, additional competition can reduce forecast accuracy and increase noise. This is particularly relevant in light of recent reversals of the Settlement and MiFID II's unbundling rules, both of which were motivated in part by promoting analyst competition for less-covered small stocks. While well intentioned, these reversals may simultaneously increase both optimism and competition for high-profile stocks, which are likely already heavily covered. Increasing optimism and competition may then reduce the aggregate quality of information supplied by analysts for these stocks. Such an increase in forecast distortions may be particularly harmful for retail investors, who are more likely to take analysts' forecasts as their honest beliefs. Future regulation would benefit from a more targeted approach to address analysts' conflict of interest.

References

Athey, S. and P. A. Haile (2007). Nonparametric approaches to auctions. In *Handbook of Econometrics*, Volume 6A, pp. 3847–3965. Elsevier.

Bajari, P., C. L. Benkard, and J. Levin (2007). Estimating dynamic models of imperfect competition. *Econometrica 75*(5), 1331–1370.

Bar-Isaac, H. and A. Gavazza (2015). Brokers' contractual arrangements in the Manhattan residential rental market. *Journal of Urban Economics 86*, 73–82.

Becker, B. and T. Milbourn (2011). How did increased competition affect credit ratings? *Journal of Financial Economics 101*(3), 493–514.

Bergemann, D. and M. Ottaviani (2021). Information markets and nonmarkets. In *Handbook of Industrial Organization*, Volume 4, pp. 593–672. Elsevier.

Bhattacharya, V., G. Illanes, and M. Padi (2025). Fiduciary duty and the market for financial advice. *Econometrica 93*(4), 1449–1480.

Bolton, P., X. Freixas, and J. Shapiro (2007). Conflicts of interest, information provision, and competition in the financial services industry. *Journal of Financial Economics 85*(2), 297–330.

Bordalo, P., N. Gennaioli, R. La Porta, and A. Shleifer (2019). Diagnostic expectations and stock returns. *The Journal of Finance 74*(6), 2839–2874.

Bordalo, P., N. Gennaioli, R. La Porta, and A. Shleifer (2024). Belief overreaction and stock market puzzles. *Journal of Political Economy 132*(5), 1450–1484.

Bouchaud, J.-P., P. Krüger, A. Landier, and D. Thesmar (2019). Sticky expectations and the profitability anomaly. *The Journal of Finance 74*(2), 639–674.

Cagé, J. (2020). Media competition, information provision and political participation: Evidence from French local newspapers and elections, 1944–2014. *Journal of Public Economics 185*, 104077.

Camara, F. (2015). Preemption and forecast accuracy: A structural approach. Working Paper, University of Southern California.

Chevalier, J. and G. Ellison (1997). Risk taking by mutual funds as a response to incentives. *Journal of Political Economy 105*(6), 1167–1200.

Chevalier, J. and G. Ellison (1999). Career concerns of mutual fund managers. *The Quarterly Journal of Economics 114*(2), 389–432.

Chu, C. S. and M. Rysman (2019). Competition and strategic incentives in the market for credit ratings: Empirics of the financial crisis of 2007. *American Economic Review 109*(10), 3514–3555.

Clark, R., J.-F. Houde, and J. Kastl (2021). The industrial organization of financial markets. In *Handbook of Industrial Organization*, Volume 5, pp. 427–520. Elsevier.

Clarke, J. and A. Subramanian (2006). Dynamic forecasting behavior by analysts: Theory and evidence. *Journal of Financial Economics 80*(1), 81–113.

Coibion, O. and Y. Gorodnichenko (2012). What can survey forecasts tell us about information rigidities? *Journal of Political Economy 120*(1), 116–159.

Compiani, G., P. A. Haile, and M. Sant'Anna (2020). Common values, unobserved heterogeneity, and endogenous entry in US offshore oil lease auctions. *Journal of Political Economy 128*(10), 3872–3912.

Deb, R., M. M. Pai, and M. Said (2018). Evaluating strategic forecasters. *American Economic Review 108*(10), 3057–3103.

Dewatripont, M. and J. Tirole (1999). Advocates. *Journal of Political Economy 107*(1), 1–39.

Egan, M. (2019). Brokers versus retail investors: Conflicting interests and dominated products. *The Journal of Finance 74*(3), 1217–1260.

Egan, M., G. Matvos, and A. Seru (2025). Arbitration with uninformed consumers. *The Review of Economic Studies 92*(6), 3888–3923.

Gentzkow, M. and J. M. Shapiro (2008). Competition and truth in the market for news. *Journal of Economic Perspectives 22*(2), 133–154.

Groysberg, B., P. M. Healy, and D. A. Maber (2011). What drives sell-side analyst compensation at high-status investment banks? *Journal of Accounting Research 49*(4), 969–1000.

Guo, Y. and L. Mota (2021). Should information be sold separately? Evidence from MiFID II. *Journal of Financial Economics 142*(1), 97–126.

Haile, P. A., H. Hong, and M. Shum (2003). Nonparametric tests for common values at first-price sealed-bid auctions. NBER Working Paper 10105, National Bureau of Economic Research.

Hendricks, K., J. Pinkse, and R. H. Porter (2003). Empirical implications of equilibrium bidding in first-price, symmetric, common value auctions. *The Review of Economic Studies 70*(1), 115–145.

Hong, H. and M. Kacperczyk (2010). Competition and bias. *The Quarterly Journal of Economics 125*(4), 1683–1725.

Hong, H. and J. D. Kubik (2003). Analyzing the analysts: Career concerns and biased earnings forecasts. *The Journal of Finance 58*(1), 313–351.

Jackson, A. R. (2005). Trade generation, reputation, and sell-side analysts. *The Journal of Finance 60*(2), 673–717.

Judd, K. L. (2020). Computational Economics Course 2020. GitHub repository, `https://kennethjudd.github.io/CompEcon2020/`. Accessed August 27, 2026.

Kadan, O., L. Madureira, R. Wang, and T. Zach (2009). Conflicts of interest and stock recommendations: The effects of the Global Settlement and related regulations. *The Review of Financial Studies 22*(10), 4189–4217.

Krigman, L., W. H. Shaw, and K. L. Womack (2001). Why do firms switch underwriters? *Journal of Financial Economics 60*(2–3), 245–284.

Lourie, B., D. Shanthikumar, and I. S. Yoo (2023). MiFID II and the unbundling of analyst research from trading execution. *Contemporary Accounting Research 40*(4), 2340–2372.

Maber, D. A., B. Groysberg, and P. M. Healy (2021). An empirical examination of sell-side brokerage analysts' published research, concierge services, and high-touch services. *European Accounting Review 30*(4), 827–853.

Marinovic, I., M. Ottaviani, and P. Sørensen (2013). Forecasters' objectives and strategies. In *Handbook of Economic Forecasting*, Volume 2, pp. 690–720. Elsevier.

Merkley, K., R. Michaely, and J. Pacelli (2017). Does the scope of the sell-side analyst industry matter? An examination of bias, accuracy, and information content of analyst reports. *The Journal of Finance 72*(3), 1285–1334.

Michaely, R. and K. L. Womack (1999). Conflict of interest and the credibility of underwriter analyst recommendations. *The Review of Financial Studies 12*(4), 653–686.

Milgrom, P. and R. J. Weber (1985). Distributional strategies for games with incomplete information. *Mathematics of Operations Research 10*(4), 619–632.

Nalebuff, B. J. and J. E. Stiglitz (1983). Prizes and incentives: Towards a general theory of compensation and competition. *The Bell Journal of Economics 14*(1), 21–43.

Ng, S. and P. Perron (2001). Lag length selection and the construction of unit root tests with good size and power. *Econometrica 69*(6), 1519–1554.

Ottaviani, M. and P. N. Sørensen (2005). Forecasting and rank-order contests. Working Paper, University of Copenhagen.

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

Pesendorfer, W. and A. Wolinsky (2003). Second opinions and price competition: Inefficiency in the market for expert advice. *The Review of Economic Studies 70*(2), 417–437.

Robles-Garcia, C. (2020). Competition and incentives in mortgage markets: The role of brokers. Working Paper, Stanford University.

Online Appendix for “Does Competition Improve Information Quality: Evidence from the Security Analyst Market”

Chuqing Jin

Toulouse School of Economics

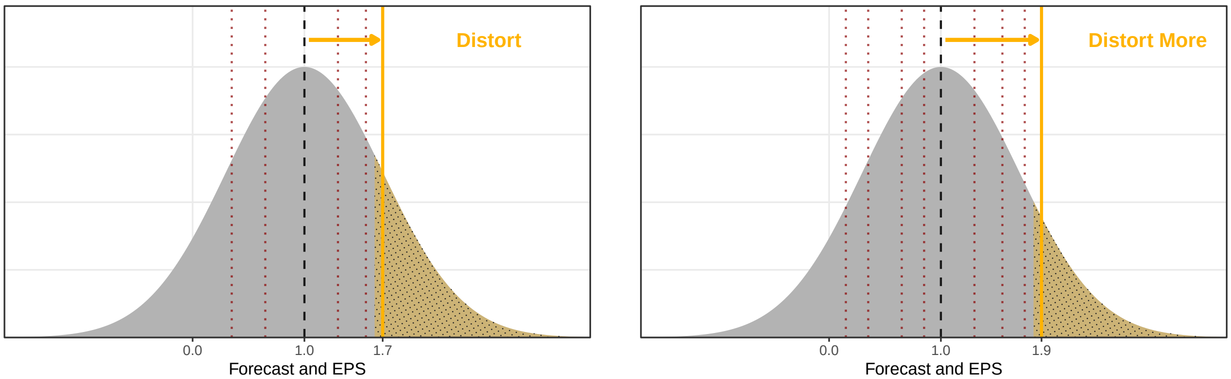
September 2026

Table of Contents

A Supplementary Tables and Figures

Figure A.1: Effect of Competition on Forecast Distortion

(a) Four rivals (b) Seven rivals



Note: Figure illustrates how the level of competition affects an analyst's forecast distortion. The red dotted vertical lines represent her rivals' forecasts, which are assumed to be known to the analyst for this illustration only. The gray unshaded area represents the analyst's posterior distribution of the earnings. The black dashed line marks the analyst's posterior mean, i.e., her *honest forecast*. The yellow pattern-shaded areas represent the probabilities of being the most accurate under the distorted forecasts. Numerical values are illustrative. In Panel a, the analyst faces four rivals. The yellow solid line is her *distorted forecast*; the yellow arrow marks the direction of the distortion. The dotted pattern on the density plot marks the earnings realizations for which the distorted forecast would be the most accurate. In Panel b, the analyst faces seven rivals. The yellow solid line is her *distorted forecast*, which lies farther from her honest forecast than in Panel a; the yellow arrow marks the direction of the larger distortion. The dotted pattern marks the earnings realizations for which this forecast would be the most accurate.

Table A.1: Summary Statistics: Main Sample

	Mean	Pctl(25)	Median	Pctl(75)
Panel A: Characteristics of Analysts				
Number of Securities Covered	7	3	6	10
Number of Industries Covered	2	1	1	2
High Prestige	0.506	0.000	1.000	1.000
Number of Lifetime Brokerage Houses	2	1	1	2
Number of Analyst-Year Observations		67,442		
Panel B: Characteristics of Brokerage Houses				
Number of Securities	77	9	27	84
Number of Industries	5	3	5	8
Number of Analysts	15	2	7	17
Number of Analysts per Security	1	1	1	1
Number of Brokerage-Year Observations		6,339		
Panel C: Characteristics of Securities				
$\log($Market Cap$)$	14.409	13.258	14.318	15.454
$\log($Book Value$)$	6.671	5.625	6.578	7.582
Profitability	0.274	0.151	0.232	0.320
Volatility $(\times 1{,}000)$	0.655	0.233	0.407	0.762
Average Monthly Return	0.014	-0.002	0.013	0.028
SP500	0.362	0.000	0.000	1.000
Share of High-Prestige Analysts	0.487	0.333	0.500	0.643
EPS Report Delay	37	26	34	45
Number of Security-Year Observations		14,528		

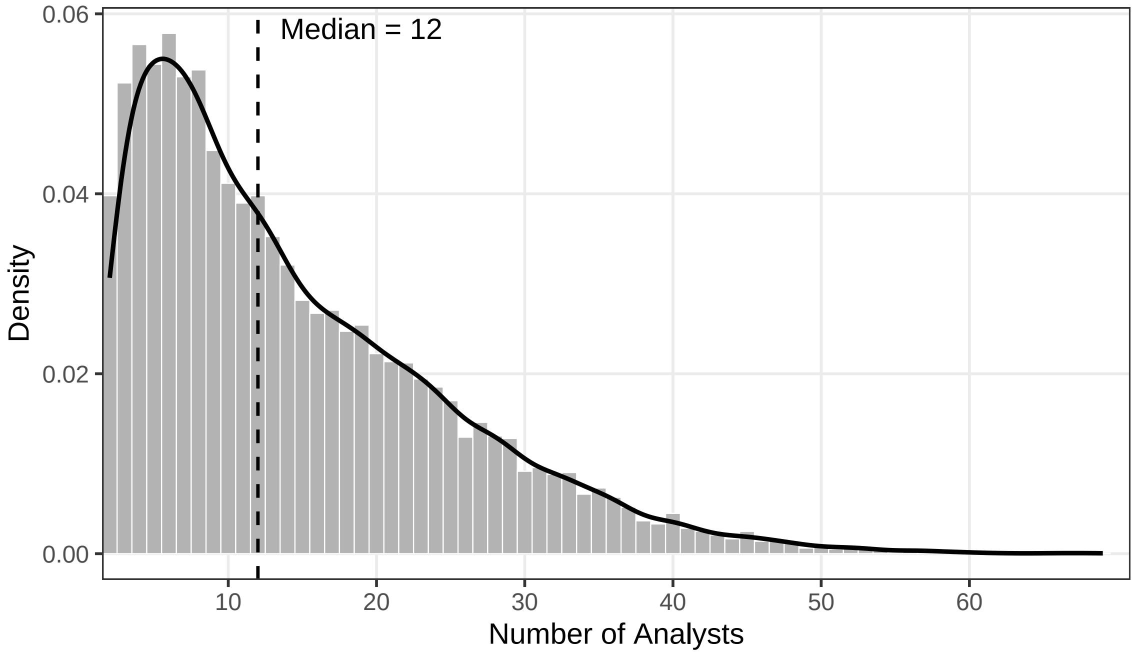
Note: Table reports summary statistics for the main analysis sample, which restricts the full analysis sample to securities observed in the data for at least 20 years. Panel A is measured at the analyst-year level, except that the number of lifetime brokerage houses is measured once per analyst over her observed career in IBES. Panel B is measured at the brokerage-year level. Panel C is measured at the security-year level. For a given security, *log(Market Cap)* is the natural logarithm of the security's fiscal-year-end market capitalization, calculated as the absolute CRSP month-end stock price multiplied by shares outstanding ($prc \times shrout$). *log(Book Value)* is the natural logarithm of the security's fiscal-year-end common book equity, calculated as Compustat total assets minus total liabilities and preferred stock ($at - lt - pstk$). *Profitability* is operating income after depreciation divided by common book equity ($oiadp/(at - lt - pstk)$). *Volatility* $(\times 1,000)$ is the variance of daily CRSP stock returns over the 12 months ending in the fiscal-year-end month and is multiplied by 1,000 for readability. *Average Monthly Return* is the average CRSP monthly stock return over the same 12-month period. *SP500* is an indicator equal to one if the security is included in the S&P 500 on its fiscal-year-end date. *Share of High-Prestige Analysts* is the fraction of analysts covering the security in a fiscal year who are employed by high-prestige brokerage houses. A brokerage house is classified as high prestige if the number of analysts it employs is at or above the 90th percentile of brokerage houses in that year. *EPS Report Delay* is the number of days between the fiscal-year-end date and the earnings announcement date for that fiscal year. Industries are defined using the first two digits of the Thomson Reuters Business Classification (TRBC) permanent industry code; the sample spans 10 industries.

Table A.2: Effect of Competition on Analysts' Forecasts: Full Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Forecast Error							
$\log(N_{jt})$	0.046	0.126	0.236	0.153	0.059	0.014	0.052
	(0.006)	(0.012)	(0.071)	(0.019)	(0.018)	(0.004)	(0.006)
Panel B: Forecast							
$\log(N_{jt}) \times Z_{jt}$	0.070	0.074	0.132	0.124	0.090	0.057	0.059
	(0.007)	(0.007)	(0.017)	(0.012)	(0.011)	(0.005)	(0.005)
Z_{jt}	0.665	0.652	0.481	0.508	0.609	0.742	0.736
	(0.021)	(0.023)	(0.051)	(0.036)	(0.034)	(0.016)	(0.017)
$\log(N_{jt})$	-0.055	-0.047	0.211	-0.102	-0.128	0.003	0.015
	(0.008)	(0.014)	(0.081)	(0.020)	(0.021)	(0.005)	(0.007)
Year FE	Y	Y	Y	Y	Y	Y	Y
Security FE						Y	Y
Security-analyst FE	Y	Y	Y	Y	Y		
Characteristics	Y	Y	Y	Y	Y		
Timing Controls	Y	Y	Y	Y	Y		
Lagged Competition IV		Y					Y
Merger IV			Y				
Early-coverage IV				Y			
Early-coverage-exit IV					Y		
Observations	425,781	425,781	425,781	321,499	319,097	466,686	466,686

Note: Table repeats the main reduced-form specifications in the full analysis sample before standardization, without restricting the sample to securities observed in the data for at least 20 years and covered by at least two analysts. Panel A uses individual analysts' forecast errors as the dependent variable. Panel B uses individual analysts' forecasts as the dependent variable. Characteristics controls include security characteristics (market capitalization, book value, profitability, return volatility, average monthly returns, an indicator for inclusion in the S&P 500 index, and the share of covering analysts from high-prestige brokerage houses), analyst characteristics (experience with the security, with the industry, and with the profession, and prestige), and forecast-timing controls (forecast order and earnings-report delay). The lagged-competition IV is the one-year lag of the log number of analysts. The merger IV consists of 15 difference-in-differences-style indicators, one for each brokerage-house merger in the sample, equal to one if and only if the security is covered by both merging brokerages, for fiscal years in or after the year of the merger. The early-coverage IV is, for each security-year, the log count of analysts who covered the security during their own first three years in the IBES sample and had entered the sample by that year; the early-coverage-exit IV additionally drops analysts after their final year in the sample. Both early-coverage columns exclude forecasts from analysts' first three years. The interaction term is instrumented with the interaction of the IV and the realized earnings. Standard errors are heteroskedasticity-consistent and clustered at the security-year level.

Figure A.2: Distribution of Competition Across Security-Years: Main Sample



Note: Figure presents the distribution of competition across security-years in the main sample, where competition is measured by the number of analysts covering a security in a year. The bars form a density histogram with one-analyst-wide bins, and the solid line plots the corresponding kernel density estimate. The dashed vertical line marks the median.

Table A.3: First-Stage Strength

	(1) Lagged Competition	(2) Merger	(3) Early Coverage	(4) Early Coverage Exit
Panel A: Forecast Error Specification				
Kleibergen–Paap rk Wald F	1712.6	13.8	1287.0	2880.6
Observations	195,892	195,892	143,092	142,863
Panel B: Forecast Specification				
Kleibergen–Paap rk Wald F	830.0	8.4	637.1	1433.0
Observations	195,892	195,892	143,092	142,863

Note: Table reports first-stage strength for Columns (2)–(5) of the main reduced-form table. Panel A corresponds to the forecast-error specification, in which $\log(N_{jt})$ is the endogenous regressor. Panel B corresponds to the forecast specification, in which $\log(N_{jt})$ and $\log(N_{jt}) \times Z_{jt}$ are endogenous. The reported statistics are heteroskedasticity- and security-year-cluster-robust Kleibergen–Paap rk Wald F-statistics. Column (1) uses $\log(N_{jt-1})$ as the excluded instrument. Column (2) uses the 15 brokerage-merger indicators jointly. Columns (3) and (4) use the early-coverage and early-coverage-exit instruments, respectively. All specifications include security-analyst and year fixed effects, security and analyst characteristics, and forecast-timing controls. The early-coverage specifications exclude forecasts from analysts' first three years.

B Data Cleaning Details

B.1 Outlier Removal and Sample Definition

To remove potential coding errors, I exclude a security if any of its forecasts has a forecast error greater than $10 or if its mean price-normalized forecast error in any year exceeds the 99.5th percentile across security-years. I also remove a small number of securities with ambiguous fiscal-period-end matches and require nonmissing analyst characteristics, security characteristics, forecast-timing controls, and analyst coverage in the immediately preceding fiscal year. This procedure results in 466,812 forecasts covering 4,709 securities, 15,953 analysts, and 1,011 brokerage houses.

B.2 Standardizing Analysts' Forecast Problems

For a given security, denote the Markov process of its realized earnings as

$$Z_t = F(Z_{t-1}) + \xi_t,$$

where $\xi_t \sim N(0, \sigma^2)$. I assume that analysts covering a security know the Markov process $F(\cdot)$ governing its realized earnings. The forecasting problem in each period is therefore equivalent to forecasting ξ_t/σ. I standardize analyst i's forecast X_{it} and realized earnings Z_t using the estimated time-series process $\hat{F}(\cdot)$ as follows:

$$\widetilde{Z}_t = \frac{Z_t - \hat{F}(Z_{t-1})}{\hat{\sigma}},$$
$$\widetilde{X}_{it} = \frac{X_{it} - \hat{F}(Z_{t-1})}{\hat{\sigma}}.$$

In practice, I first run a unit-root test on the realized earnings of each security using the MZ test with GLS-detrended data (Ng and Perron, 2001). If I cannot reject a unit root, I fit a unit-root process for $\hat{F}(\cdot)$; otherwise, I fit an AR(1) process with a time trend. For each security-year, I then standardize realized earnings and forecasts by subtracting the fitted value of realized earnings and dividing by the estimated standard deviation of the time-series shock. For simplicity, I subsequently use X_{it} and Z_t to denote the standardized forecast and realized earnings.

C Proofs

C.1 Proof of Proposition 1

I follow Proposition 3 of Ottaviani and Sørensen (2005) and incorporate absolute accuracy and optimism in the payoff function.

Proof. Consider an arbitrary N. I apply results from Milgrom and Weber (1985) (henceforth MW) and show that conditions (i)–(vi) of their Theorem 4 are satisfied. $\mathcal{Z}$ in my model corresponds to both the metric space of the environmental variable T_0 and the action space A_i in MW. Assumption 1 ensures the players' types are independent conditional on the environmental variable, so (i) is satisfied. Analysts' signals are continuously distributed, so (ii) is satisfied. An analyst's payoff function depends only on the realized earnings Z and the vector of forecasts X so (iii) is satisfied. Finally, because $\mathcal{Z}$ is finite, analysts' payoffs are equicontinuous by Proposition 1(a) of MW and the metric space of the environmental variable is also finite, i.e., (iv), (v), and (vi) are satisfied. By Theorem 4 of MW, there exists an equilibrium in pure strategies. ∎

C.2 Extension of Proposition 1 to Allow Common Signals

Now suppose analysts' signals are correlated through a common signal $C \in \mathcal{C} \subset \mathbb{R}$ where $\mathcal{C}$ is finite. Specifically, let $S_i = (C, \tilde{S}_i)$ where the common signal C is observed by all analysts in the contest and $\tilde{S}_i$ is analyst i's private signal. In this case, Assumption 1 is no longer satisfied. To guarantee that an equilibrium exists, it is sufficient to assume that analysts' private signals are uncorrelated conditional on the realized earnings and the common signal:

Assumption 3. $\tilde{S}_i \perp\!\!\!\perp \tilde{S}_j | C, Z$ for all $j \neq i$.

Corollary 1. If Assumption 3 holds, analysts' private signals $\tilde{S}_i$ are continuously distributed, and $\mathcal{C}$ and $\mathcal{Z}$ are finite, there exists a pure-strategy equilibrium.

Proof. Now (C, Z) corresponds to the environmental variable T_0 in MW and $\mathcal{C} \times \mathcal{Z}$ corresponds to its metric space. The rest of the argument follows the proof for Proposition 1. ∎

Intuitively, given Assumption 3, the model with a common signal is equivalent to a two-stage game. In the first stage, analysts receive a common signal and update their common prior for the realized earnings. In the second stage, analysts receive private signals and choose forecasts. The second stage is the same as the model described in Section 4. The common signal only alters analysts' common prior for the realized earnings.

C.3 Proof of Proposition 2

Proof. The assumed signal structure implies $S_i|Z \sim N(Z, 1/\tau)$; the normalization $Z \sim N(0,1)$ is without loss of generality. Denote the honest forecast strategy, which is analysts' posterior mean conditional on receiving signal s, by

$$\beta_0(s) \equiv \mathbb{E}[Z \mid S_i = s] = as, \qquad a \equiv \frac{\tau}{1+\tau}.$$

The posterior distribution is

$$Z \mid S_i = s \sim N\left(as, \frac{1}{1+\tau}\right).$$

Let ϕ and Φ denote the standard normal density and distribution function. Let $p_0(X, Z)$ denote the probability that analyst i beats one arbitrary rival when all rivals use the honest forecast strategy β_0. Let $\Xi_0(X_i, Z, N, k)$ denote the marginal probability of beating exactly a given set of $N-k$ rivals when all players use the honest forecast strategy β_0.

The proof strategy is to show that when rivals use the honest forecast strategy, the first-order condition of an analyst is positive at the honest forecast for some positive signal $s \in (0, \bar{S}_N)$, and negative at the honest forecast for some negative signal $s \in (-\bar{S}_N, 0)$. This would imply that there exists a deviation from the honest forecast in the direction of the signal that increases the analyst's expected payoff.

Step 1: Simplify p_0 and Ξ_0 given honest forecasts and normality Let t denote the distance between the realized earnings and the forecast, i.e., $t \equiv |Z - X_i|$. When $Z > X_i = \beta_0(s) = as$, we can write $Z = as + t$, where $t > 0$. Then, rival j beats analyst i if and only if $|X_j - Z| < |X_i - Z|$, or equivalently

$$as < X_j < 2Z - as = as + 2t.$$

Because rivals report $X_j = \beta_0(S_j) = aS_j$, the corresponding signal interval is

$$s < S_j < s + \frac{2t}{a}.$$

For ease of exposition, denote

$$c \equiv \frac{2-a}{a} = \frac{2+\tau}{\tau} > 1.$$

Because $S_j \mid Z \sim N(Z, 1/\tau)$, we have

$$p_0(as, as+t) = 1 - \left[\Phi\left(\sqrt{\tau}\left((1-a)s + ct\right)\right) - \Phi\left(\sqrt{\tau}\left((1-a)s - t\right)\right)\right]. \tag{12}$$

Similarly, when $Z < X_i = \beta_0(s) = as$, we can write $Z = as - t$, where $t > 0$. Then, rival j beats i if and only if

$$as - 2t < X_j < as,$$

or equivalently

$$s - \frac{2t}{a} < S_j < s.$$

Therefore,

$$p_0(as, as-t) = 1 - \left[\Phi\left(\sqrt{\tau}\left((1-a)s + t\right)\right) - \Phi\left(\sqrt{\tau}\left((1-a)s - ct\right)\right)\right]. \tag{13}$$

The marginal probability of beating $N-1$ rivals and being ranked first is given by

$$\begin{aligned}\Xi_0(X_i, Z, N, 1) =& \big[\mathbb{1}(X_i \le Z) - \mathbb{1}(X_i > Z)\big] \\ & \big[\frac{g(\beta_0^{-1}(X_i)|Z)}{\beta_0'(\beta_0^{-1}(X_i))} + \frac{g(\beta_0^{-1}(2Z - X_i)|Z)}{\beta_0'(\beta_0^{-1}(2Z - X_i))}\big](N-1)p_0(X_i, Z)^{N-2}.\end{aligned} \tag{14}$$

Define the object in the second bracket as $B^0(X, Z)$, which is the marginal probability of beating a given rival when all rivals use the honest forecast strategy β_0:

$$B^0(X, Z) \equiv \frac{g\big(\beta_0^{-1}(X) \mid Z\big)}{\beta_0'\big(\beta_0^{-1}(X)\big)} + \frac{g\big(\beta_0^{-1}(2Z - X) \mid Z\big)}{\beta_0'\big(\beta_0^{-1}(2Z - X)\big)}.$$

Because $g(u \mid Z) = \sqrt{\tau}\,\phi\left(\sqrt{\tau}(u - Z)\right)$, when $Z = as + t > X_i = as$ for $t > 0$, we obtain

$$B^0(as, as+t) = \frac{\sqrt{\tau}}{a}\left[\phi\left(\sqrt{\tau}\left((1-a)s - t\right)\right) + \phi\left(\sqrt{\tau}\left((1-a)s + ct\right)\right)\right]. \tag{15}$$

When $Z = as - t < X_i = as$ for $t > 0$,

$$B^0(as, as-t) = \frac{\sqrt{\tau}}{a}\left[\phi\left(\sqrt{\tau}\left((1-a)s + t\right)\right) + \phi\left(\sqrt{\tau}\left((1-a)s - ct\right)\right)\right]. \tag{16}$$

Now we derive the distribution of t. When $Z > X_i$, $t = Z - as$, so $t \mid S_i = s \sim N\left(0, \frac{1}{1+\tau}\right)$, with density $\psi(t) = \sqrt{\frac{1+\tau}{2\pi}} \exp\left(-\frac{1+\tau}{2}t^2\right)$. When $Z < X_i$, we can derive the same expression for $\psi(t)$.

In a winner-takes-all contest without optimism, only the $k = 1$ term of the relative-accuracy component remains, and the absolute-accuracy term $-2[X_i - \mathbb{E}(Z|S_i)]$ vanishes at the honest forecast $X_i = \beta_0(s)$. Putting these together and letting $n = N - 2$, the first-order condition at the honest forecast is

$$M_N(s) \equiv \gamma_1(N) \int_{-\infty}^{\infty} \Xi_0\big(\beta_0(s), Z, N, 1\big) f(Z \mid s) dZ \tag{17}$$

$$= \gamma_1(N)(N-1) \int_0^{\infty} \Big[B^0(as, as+t)p_0(as, as+t)^n - B^0(as, as-t)p_0(as, as-t)^n\Big]\psi(t)dt. \tag{18}$$

Therefore, the goal is to find the sign of $M_N(s)$ when s is around 0. By symmetry of the standard normal distribution, when $s = 0$, $B^0(0, t) = B^0(0, -t)$ and $p_0(0, t) = p_0(0, -t)$, so $M_N(0) = 0$. To find signs of $M_N(s)$ in the neighborhood of $s = 0$, we then consider the derivative of $M_N(s)$ at zero.

Step 2: Compute derivative at zero signal Let the integrand inside brackets in (18) be denoted by $H(s, t)$:

$$H(s, t) \equiv B^0(as, as+t)p_0(as, as+t)^n - B^0(as, as-t)p_0(as, as-t)^n.$$

Then

$$M_N(s) = \gamma_1(N)(N-1) \int_0^{\infty} H(s, t)\psi(t)dt.$$

By the Leibniz rule, we have,

$$M_N'(0) = \gamma_1(N)(N-1) \int_0^{\infty} H_s(0, t)\psi(t)dt.$$

We will do a change of variable to simplify the expression. Let $x = \sqrt{\tau}t$. At $s = 0$, equations (12) and (13) give

$$p_0(0, t) = p_0(0, -t) = 2 - \Phi(x) - \Phi(cx).$$

Similarly, equations (15) and (16) give

$$B^0(0, t) = B^0(0, -t) = \frac{\sqrt{\tau}}{a}\{\phi(x) + \phi(cx)\}. \tag{19}$$

Also, $dt = dx/\sqrt{\tau}$ and $\frac{1+\tau}{2\tau} = \frac{c+1}{4}$, so

$$\begin{aligned}\psi(t)dt &= \sqrt{\frac{1+\tau}{2\pi}} \exp\left(-\frac{1+\tau}{2}t^2\right) dt \\ &= \sqrt{\frac{1+\tau}{2\pi\tau}} \exp\left(-\frac{1+\tau}{2\tau}x^2\right) dx \\ &= \sqrt{\frac{1+\tau}{2\pi\tau}} \exp\left(-\frac{c+1}{4}x^2\right) dx \end{aligned} \tag{20}$$

To compute $H_s(0,t)$, we differentiate its components with respect to s, starting with the probability $p_0(X,Z)$. From equation (12),

$$\frac{\partial}{\partial s} p_0(as, as+t) = -\sqrt{\tau}(1-a)\phi\left(\sqrt{\tau}((1-a)s+ct)\right) + \sqrt{\tau}(1-a)\phi\left(\sqrt{\tau}((1-a)s-t)\right).$$

Thus, at $s=0$,

$$\left.\frac{\partial}{\partial s} p_0(as, as+t)\right|_{s=0} = \sqrt{\tau}(1-a)\{\phi(x) - \phi(cx)\}.$$

Similarly, from equation (13),

$$\left.\frac{\partial}{\partial s} p_0(as, as-t)\right|_{s=0} = \sqrt{\tau}(1-a)\{\phi(cx) - \phi(x)\}.$$

Therefore,

$$\left.\frac{\partial}{\partial s} p_0(as, as+t)\right|_{s=0} - \left.\frac{\partial}{\partial s} p_0(as, as-t)\right|_{s=0} = 2\sqrt{\tau}(1-a)\{\phi(x) - \phi(cx)\}. \tag{21}$$

Next differentiate the term B^0. Since $\phi'(u) = -u\phi(u)$,

$$\frac{\partial}{\partial s} B^0(as, as+t) = \frac{\tau(1-a)}{a}\left[\phi'\left(\sqrt{\tau}((1-a)s-t)\right) + \phi'\left(\sqrt{\tau}((1-a)s+ct)\right)\right].$$

At $s=0$, this becomes

$$\left.\frac{\partial}{\partial s} B^0(as, as+t)\right|_{s=0} = \frac{\tau(1-a)}{a} x\{\phi(x) - c\phi(cx)\}.$$

Analogously,

$$\left.\frac{\partial}{\partial s} B^0(as, as-t)\right|_{s=0} = -\frac{\tau(1-a)}{a} x\{\phi(x) - c\phi(cx)\}.$$

Therefore,

$$\frac{\partial}{\partial s}B^0(as, as+t)\bigg|_{s=0} - \frac{\partial}{\partial s}B^0(as, as-t)\bigg|_{s=0} = 2\frac{\tau(1-a)}{a}x\{\phi(x) - c\phi(cx)\}. \tag{22}$$

Then, using the product rule on $H(s,t)$, equations (19), (21) and (22) imply

$$H_s(0,t) = 2\frac{\tau(1-a)}{a}\Big[p_0(x)^n x\{\phi(x) - c\phi(cx)\} + np_0(x)^{n-1}\{\phi(x) + \phi(cx)\}\{\phi(x) - \phi(cx)\}\Big],$$

where $p_0(x)$ is defined as

$$p_0(x) \equiv p_0(0,t) = p_0(0,-t) = 2 - \Phi(x) - \Phi(cx). \tag{23}$$

When $n = 0$, i.e., $N = 2$, the second term in the bracket will be omitted.

As $a = \tau/(1+\tau)$, $\frac{\tau(1-a)}{a} = 1$. Hence

$$H_s(0,t) = 2\Big[p_0(x)^n x\{\phi(x) - c\phi(cx)\} + np_0(x)^{n-1}\{\phi(x) + \phi(cx)\}\{\phi(x) - \phi(cx)\}\Big]. \tag{24}$$

Finally, equations (20) and (24) yield

$$M_N'(0) = 2\gamma_1(N)(N-1)\sqrt{\frac{1+\tau}{2\pi\tau}}\, I_n(c),$$

where

$$\begin{aligned} I_n(c) = \int_0^\infty \exp(-\frac{c+1}{4}x^2)\Big[&p_0(x)^n x\{\phi(x) - c\phi(cx)\} \\ &+ np_0(x)^{n-1}\{\phi(x) + \phi(cx)\}\{\phi(x) - \phi(cx)\}\Big]dx, \end{aligned} \tag{25}$$

with the second term omitted when $n = 0$.

Note that the sign of $M_N'(0)$ is the same as the sign of $I_n(c)$, because the rest of the expression is positive. We will show that $I_n(c) > 0$ for every integer $n \geq 0$ and every $c > 1$ in the case with two analysts ($N = 2$) and with at least three analysts ($N \geq 3$).

Step 3: Show $I_n(c)$ is positive

Case 1: $N = 2$ This implies that $n = 0$, so using the pdf of the standard normal distribution,

$$\begin{aligned} I_0(c) &= \int_0^\infty e^{-\frac{c+1}{4}x^2} x\{\phi(x) - c\phi(cx)\}dx \\ &= \frac{1}{\sqrt{2\pi}} \int_0^\infty x \left[e^{-\frac{c+1}{4}x^2} e^{-x^2/2} - ce^{-\frac{c+1}{4}x^2} e^{-c^2x^2/2}\right] dx \\ &= \frac{1}{\sqrt{2\pi}} \left[\int_0^\infty xe^{-\frac{c+3}{4}x^2} dx - c\int_0^\infty xe^{-\frac{2c^2+c+1}{4}x^2} dx\right]. \end{aligned}$$

For any $\lambda > 0$,

$$\int_0^\infty xe^{-\lambda x^2} dx = \frac{1}{2\lambda}.$$

Therefore,

$$\begin{aligned} I_0(c) &= \frac{1}{\sqrt{2\pi}} \left[\frac{2}{c+3} - c\frac{2}{2c^2+c+1}\right] \\ &= \frac{2}{\sqrt{2\pi}} \left[\frac{1}{c+3} - \frac{c}{2c^2+c+1}\right] \\ &= \frac{2}{\sqrt{2\pi}} \frac{2c^2+c+1-c(c+3)}{(c+3)(2c^2+c+1)} \\ &= \frac{2(c-1)^2}{\sqrt{2\pi}(c+3)(2c^2+c+1)}. \end{aligned}$$

Since $c > 1$, this expression is strictly positive.

Case 2: $N \geq 3$ This implies that $n \geq 1$. I will integrate by parts to simplify $I_n(c)$, which I restate here for convenience,

$$\begin{aligned} I_n(c) = \int_0^\infty \exp(-\frac{c+1}{4}x^2)\Big[&p_0(x)^n x\{\phi(x) - c\phi(cx)\} \\ &+ np_0(x)^{n-1}\{\phi(x) + \phi(cx)\}\{\phi(x) - \phi(cx)\}\Big]dx. \end{aligned} \tag{26}$$

The second term in the integrand of $I_n(c)$ is

$$np_0(x)^{n-1}\{\phi(x) + \phi(cx)\}\{\phi(x) - \phi(cx)\} = np_0(x)^{n-1}\{\phi(x)^2 - \phi(cx)^2\}.$$

Note that

$$(p_0(x)^n)' = -np_0(x)^{n-1}\{\phi(x) + c\phi(cx)\}.$$

Therefore,

$$np_0(x)^{n-1}\{\phi(x)^2 - \phi(cx)^2\} = -(p_0(x)^n)' \frac{\phi(x)^2 - \phi(cx)^2}{\phi(x) + c\phi(cx)}.$$

Define

$$R(x) = \frac{\phi(x)^2 - \phi(cx)^2}{\phi(x) + c\phi(cx)}. \tag{27}$$

Therefore, the second-term integral is

$$-\int_0^\infty (p_0(x)^n)' \exp(-\frac{c+1}{4}x^2)R(x)dx.$$

Integrating by parts gives

$$\begin{aligned} &-\int_0^\infty (p_0(x)^n)' \exp(-\frac{c+1}{4}x^2)R(x)dx \\ = &-\left[p_0(x)^n \exp(-\frac{c+1}{4}x^2)R(x)\right]_0^\infty + \int_0^\infty p_0(x)^n \frac{d}{dx}\left[\exp(-\frac{c+1}{4}x^2)R(x)\right]dx. \end{aligned}$$

At $x = 0$, $R(0) = 0$. As $x \to \infty$, because $c > 1$,

$$r = r(x) \equiv \frac{\phi(cx)}{\phi(x)} = \exp\left(-\frac{c^2-1}{2}x^2\right) \to 0.$$

Therefore,

$$R(x) = \frac{\phi(x)^2 - \phi(cx)^2}{\phi(x) + c\phi(cx)} = \phi(x)\frac{1-r^2}{1+cr} \to 0.$$

Meanwhile, $p_0(x)^n \leq 1$ and $\exp(-(c+1)x^2/4) \to 0$. Substituting these expressions back into $I_n(c)$ gives

$$I_n(c) = \int_0^\infty p_0(x)^n K_c(x)dx,$$

where

$$K_c(x) = \exp(-\frac{c+1}{4}x^2)x\{\phi(x) - c\phi(cx)\} + \frac{d}{dx}\left[\exp(-\frac{c+1}{4}x^2)R(x)\right]. \tag{28}$$

So the problem is further reduced to determining the sign of $K_c(x)$ for $x \geq 0$. Note that $r(x)$ decreases continuously from 1 to 0 as x increases from 0 to ∞.

Define simplifying notations:

$$
\begin{aligned}
A(r) \equiv& (1-r^2)/(1+cr) \\
\alpha \equiv& (c+1)/4
\end{aligned}
$$

Using $\phi'(x) = -x\phi(x)$ and $r'(x) = -(c^2-1)xr$, the derivative term in K_c is

$$
\begin{aligned}
\frac{d}{dx}\left[\exp(-\alpha x^2)R(x)\right] &= \frac{d}{dx}\left[\exp(-\alpha x^2)\phi(x)A(r)\right] \\
&= \exp(-\alpha x^2)x\phi(x)\left[-(1+2\alpha)A(r) - (c^2-1)rA'(r)\right].
\end{aligned}
$$

Because $1+2\alpha = (c+3)/2$, $\exp(-\alpha x^2)x\{\phi(x) - c\phi(cx)\} = \exp(-\alpha x^2)x\phi(x)(1-cr)$, and

$$
A'(r) = \frac{(-2r)(1+cr) - c(1-r^2)}{(1+cr)^2} = -\frac{c+2r+cr^2}{(1+cr)^2}.
$$

we write

$$
K_c(x) = \exp(-\alpha x^2)x\phi(x)B(r),
$$

where

$$
B(r) = 1 - cr - \frac{c+3}{2}\frac{1-r^2}{1+cr} + (c^2-1)r\frac{c+2r+cr^2}{(1+cr)^2}.
$$

Putting $B(r)$ over the common denominator $2(1+cr)^2$ gives

$$
\begin{aligned}
B(r) = \frac{1}{2(1+cr)^2}\Big[&2(1-cr)(1+cr)^2 - (c+3)(1-r^2)(1+cr) \\
&+ 2(c^2-1)r(c+2r+cr^2)\Big].
\end{aligned}
$$

Expanding the numerator,

$$
\begin{aligned}
2(1-cr)(1+cr)^2 &= 2 + 2cr - 2c^2r^2 - 2c^3r^3, \\
-(c+3)(1-r^2)(1+cr) &= -(c+3) - c(c+3)r + (c+3)r^2 + c(c+3)r^3, \\
2(c^2-1)r(c+2r+cr^2) &= 2c(c^2-1)r + 4(c^2-1)r^2 + 2c(c^2-1)r^3.
\end{aligned}
$$

Collecting coefficients yields

$$
B(r) = \frac{c+1}{2(1+cr)^2}\left[cr^3 + (2c-1)r^2 + (2c^2-3c)r - 1\right].
$$

Define

$$G_c(r) \equiv cr^3 + (2c-1)r^2 + (2c^2-3c)r - 1. \tag{29}$$

For $x > 0$,

$$\exp(-\alpha x^2) x\phi(x) \frac{c+1}{2(1+cr)^2} > 0.$$

Therefore, the sign of $K_c(x)$ is the same as the sign of $G_c(r(x))$.

It remains to show that G_c has exactly one zero in $(0,1)$. We have

$$G_c(0) = -1 < 0, \qquad G_c(1) = 2(c^2-1) > 0,$$

and

$$G_c'''(r) = 6cr + 4c - 2 > 0 \qquad \text{for all } r \in [0,1],\ c > 1.$$

Thus G_c' is strictly increasing. If $G_c'(0) \geq 0$, then G_c is strictly increasing on $[0,1]$ and has exactly one zero. If $G_c'(0) < 0$, then G_c first decreases and then increases, with at most one turning point; since it starts below zero at $r = 0$ and ends above zero at $r = 1$, it again has exactly one zero. Hence G_c is negative below the zero and positive above it. Now we convert the signs to the signs of $K_c(x)$. Because when x increases from 0 to ∞, $r(x)$ decreases from 1 to 0, $K_c(x)$ is first positive and then negative, with a single sign change.

Now we go back to show that for every integer $n \geq 1$ and every $c > 1$,

$$I_n(c) = \int_0^\infty p_0(x)^n K_c(x) dx > 0.$$

We will approach this with integration by parts again.

Let

$$J_c(y) = \int_0^y K_c(x) dx.$$

We first show that $J_c(y) \geq 0$ for every $y \geq 0$. Consider the limit of this function when $y \to \infty$:

$$J_c(\infty) = I_0(c) > 0.$$

To see the equality explicitly, integrate the definition of K_c:

$$J_c(y) = \int_0^y \exp(-\frac{c+1}{4}x^2) x\{\phi(x) - c\phi(cx)\} dx + \exp(-\frac{c+1}{4}y^2) R(y) - R(0).$$

Since $R(0) = 0$ and $\exp(-(c+1)y^2/4)R(y) \to 0$ as $y \to \infty$,

$$J_c(\infty) = \int_0^\infty e^{-\frac{c+1}{4}x^2} x\{\phi(x) - c\phi(cx)\}dx = I_0(c) > 0.$$

As established above, there exists $x^* > 0$ such that $K_c(x) \geq 0$ on $[0, x^*]$ and $K_c(x) \leq 0$ on $[x^*, \infty)$. Hence, for $0 \leq y \leq x^*$,

$$J_c(y) = \int_0^y K_c(x)dx \geq 0.$$

For $y \geq x^*$,

$$J_c(y) = J_c(\infty) - \int_y^\infty K_c(x)dx.$$

Since $K_c(x) \leq 0$ for $x \geq y \geq x^*$,

$$\int_y^\infty K_c(x)dx \leq 0,$$

and therefore

$$J_c(y) \geq J_c(\infty) > 0.$$

Thus

$$J_c(y) \geq 0 \qquad \text{for all } y \geq 0. \tag{30}$$

Now we perform the integration by parts. Because $J_c'(x) = K_c(x)$,

$$I_n(c) = \int_0^\infty p_0(x)^n K_c(x)dx = \int_0^\infty p_0(x)^n dJ_c(x).$$

For a finite upper bound A, integration by parts gives

$$\int_0^A p_0(x)^n dJ_c(x) = [p_0(x)^n J_c(x)]_0^A - \int_0^A J_c(x)d(p_0(x)^n).$$

Let $A \to \infty$. At $x = 0$, $p_0(0) = 1$ and $J_c(0) = 0$. At infinity, $p_0(x) \to 0$ while $J_c(x)$ converges to the finite number $I_0(c)$. Therefore,

$$I_n(c) = -\int_0^\infty J_c(x)d(p_0(x)^n). \tag{31}$$

Finally,

$$p_0'(x) = -\phi(x) - c\phi(cx) < 0.$$

Thus, for $n \geq 1$, $d(p_0(x)^n) = np_0(x)^{n-1}p_0'(x)dx < 0$. Combining this with (30) and (31) gives

$$I_n(c) > 0.$$

Step 4: Put all components together In Step 3, we show that

$$I_n(c) > 0 \qquad \text{for every integer } n \geq 0 \text{ and every } c > 1.$$

From Step 2, this implies that

$$M_N'(0) > 0.$$

By symmetry of the standard normal distribution, $M_N(0) = 0$. Hence

$$M_N(s) = M_N'(0)s + o(s).$$

There exists $\bar{S}_N > 0$ such that $M_N(s) > 0$ for all $s \in (0, \bar{S}_N)$. For such signals, the derivative of expected payoff with respect to X_i, evaluated at $X_i = \beta_0(s)$, is positive. Therefore, a small upward deviation from the posterior mean increases expected payoff. By symmetry of the game around zero, $M_N(-s) = -M_N(s)$, so for the same $\bar{S}_N$, the derivative at $X_i = \beta_0(s)$ is negative for all $s \in (-\bar{S}_N, 0)$, and a small downward deviation from the posterior mean increases expected utility.

■

C.4 Sufficient Condition for Identification

This section provides a sufficient condition for identification when the reward for any rank of relative accuracy is homogeneous across different N's. Specifically, consider $\mathcal{N} = \{2, 3, ..., N_{max}\}$, where $N_{max} \geq 4$. I impose the following constraints on the payoff parameters,

$$\gamma_k(N) = \begin{cases} \gamma_1 & \textit{if } k = 1, \\ \gamma_k & \textit{if } 2 \leq k \leq N-2 \textit{ and } N \geq 3, \\ 0 & \textit{if } k = N-1 \textit{ and } N \geq 3, \end{cases}$$

where $\gamma_k \in \mathbb{R}$ for all k. Let $F_{X|N}$ denote the distribution of equilibrium forecasts conditional on the number of analysts. Taking expectations of equation (5) with respect to $F_{X|N}$, define the coefficient on $\gamma_k(N)$ as

$$a_{N,k} \equiv \binom{N-1}{N-k} \int_{\mathbb{R}} \int_{\mathbb{R}} \Xi(X, Z, N, k) \tilde{f}(Z \mid X, N)\, dZ\, dF_{X|N}(X).$$

Proposition 3. The payoff function is globally identified if $a_{2,1} \neq a_{3,1}$ and $a_{N,N-2} \neq 0$ for all $N \geq 4$.

Proof. Taking expectations of equation (5) with respect to $F_{X|N}$, I can rewrite the first-order conditions as

$$A\gamma = 2[\mathbb{E}(X|N) - \mathbb{E}(Z|N)],$$

where

$$A = \begin{pmatrix} 1 & a_{2,1} & 0 & \dots & 0 \\ 1 & a_{3,1} & 0 & \dots & 0 \\ 1 & a_{4,1} & a_{4,2} & \dots & 0 \\ & & & \dots & \\ 1 & a_{N_{max}-1,1} & a_{N_{max}-1,2} & \dots & 0 \\ 1 & a_{N_{max},1} & a_{N_{max},2} & \dots & a_{N_{max},N_{max}-2} \end{pmatrix}, \gamma = \begin{pmatrix} \gamma_o \\ \gamma_1 \\ \gamma_2 \\ \dots \\ \gamma_{N_{max}-2} \end{pmatrix}.$$

To show that the payoff function is globally identified, it suffices to show that there is a unique solution to γ, or equivalently, A is invertible.

The matrix A is square. Subtract its first row from every subsequent row. This operation leaves the determinant unchanged and makes the first column equal to $(1, 0, \dots, 0)^T$. Expanding along that column leaves a lower-triangular matrix whose diagonal entries are

$$a_{3,1} - a_{2,1}, \quad a_{4,2}, \quad a_{5,3}, \quad \dots, \quad a_{N_{\max},N_{\max}-2}.$$

Therefore,

$$\det(A) = (a_{3,1} - a_{2,1}) \prod_{N=4}^{N_{\max}} a_{N,N-2}.$$

Under the conditions in the proposition, $\det(A) \neq 0$. Hence A is invertible, the system has a unique solution for γ, and the payoff function is globally identified.

■

The sufficient condition for identification in this proposition means that the marginal probability of being the most accurate is different between contests with two analysts and three analysts, and that the marginal probability of receiving rank $N-2$ is nonzero for $N \geq 4$.

C.5 Individual versus Consensus Forecast Errors

Lemma 1. Suppose $Z \sim N(0,1)$ and $S_i = Z + \epsilon_i$, where the ϵ_i's are independently distributed as $N(0, 1/\tau)$, independent of Z, and $\tau > 0$. Fix the number of analysts $N \geq 1$, and let $\mathcal{B}_N$ denote the set of symmetric forecast strategies $\beta_N : \mathbb{R} \to \mathbb{R}$ satisfying $\mathbb{E}|\beta_N(S_i)| < \infty$. For any $\beta_N \in \mathcal{B}_N$, let $X_i = \beta_N(S_i)$ and $X^* \equiv N^{-1} \sum_{i=1}^{N} X_i$. The ex ante expected individual and consensus forecast errors are, respectively,

$$\mathcal{L}_I(\beta_N; N) \equiv \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N} |\beta_N(S_i) - Z|\right],$$
$$\mathcal{L}_C(\beta_N; N) \equiv \mathbb{E}\left[\left|\frac{1}{N}\sum_{i=1}^{N} \beta_N(S_i) - Z\right|\right].$$

Over all strategies in $\mathcal{B}_N$, the individual forecast error is minimized by the honest forecast strategy

$$\beta_N^I(s) = a_I s, \qquad a_I = \frac{\tau}{1+\tau},$$

whereas the consensus forecast error is minimized by

$$\beta_N^C(s) = a_C s, \qquad a_C = \frac{N\tau}{1+N\tau}.$$

The corresponding minimum forecast errors are

$$\mathcal{L}_I(\beta_N^I; N) = \sqrt{\frac{2}{\pi(1+\tau)}},$$
$$\mathcal{L}_C(\beta_N^C; N) = \sqrt{\frac{2}{\pi(1+N\tau)}}.$$

For $N > 1$,

$$a_C - a_I = \frac{(N-1)\tau}{(1+\tau)(1+N\tau)} > 0.$$

Thus, the strategy that minimizes consensus forecast error puts more weight on analysts' private signals than the strategy that minimizes individual forecast error.

Proof. First, consider individual forecast error. The posterior distribution of Z conditional on analyst i's signal is

$$Z \mid S_i = s \sim N\left(\frac{\tau}{1+\tau}s, \frac{1}{1+\tau}\right).$$

Under absolute loss, the optimal forecast conditional on $S_i = s$ is the conditional median of Z. Normality implies that the conditional median equals the conditional mean, so it is $\beta_N^I(s) = \tau s/(1+\tau)$. Therefore, for any $\beta_N \in \mathcal{B}_N$ and every $s \in \mathbb{R}$,

$$\mathbb{E}\left[|\beta_N(s) - Z| \mid S_i = s\right] \geq \mathbb{E}\left[|\beta_N^I(s) - Z| \mid S_i = s\right].$$

Integrating this inequality over the distribution of S_i and averaging across analysts proves that β_N^I globally minimizes $\mathcal{L}_I$. Therefore, the forecast bias, defined as the difference between the forecast and the realized earnings $\beta_N^I(s) - Z$ is normally distributed with mean zero and variance $1/(1+\tau)$, which gives $\mathcal{L}_I(\beta_N^I; N) = \sqrt{2/[\pi(1+\tau)]}$.

Next, consider consensus forecast error. Conditional on the vector of signals $S = (S_1, \ldots, S_N)$,

$$Z \mid S \sim N\left(\frac{\tau}{1+N\tau}\sum_{i=1}^N S_i, \frac{1}{1+N\tau}\right).$$

For any $\beta_N \in \mathcal{B}_N$, the consensus X^* is a function of S. Its conditional expected absolute error is therefore minimized by the conditional median

$$m(S) \equiv \frac{\tau}{1+N\tau}\sum_{i=1}^N S_i.$$

This lower bound is attainable by the symmetric strategy $\beta_N^C(s) = N\tau s/(1+N\tau)$, because

$$\frac{1}{N}\sum_{i=1}^N \beta_N^C(S_i) = \frac{\tau}{1+N\tau}\sum_{i=1}^N S_i = m(S).$$

Hence, β_N^C globally minimizes $\mathcal{L}_C$ over $\mathcal{B}_N$. Its forecast bias is mean-zero normal with variance $1/(1+N\tau)$, which gives $\mathcal{L}_C(\beta_N^C; N) = \sqrt{2/[\pi(1+N\tau)]}$. Finally,

$$a_C - a_I = \frac{(N-1)\tau}{(1+\tau)(1+N\tau)} > 0$$

for $N > 1$. ■

C.6 Analytical Information-Quality Measures

This section derives the four information-quality measures used in the counterfactual analysis. These expressions allow me to calculate the population moments directly, without simulating contests.

Lemma 2. Suppose $Z \sim N(0,1)$ and $S_i = Z + \epsilon_i$, where the ϵ_i's are independently distributed as $N(0, 1/\tau)$, independent of Z, and $\tau > 0$. Fix a payoff specification and a number of analysts $N \geq 2$. Suppose the symmetric equilibrium strategy is affine:

$$X_i = \beta_N(S_i) = a_N S_i + b_N,$$

where a_N and b_N denote the slope and intercept, respectively. The payoff-specification index is suppressed. Let $X^* \equiv N^{-1} \sum_{i=1}^N X_i$ denote the consensus forecast, and let ϕ and Φ denote the standard normal density and distribution functions. Define

$$\sigma_{I,N}^2 \equiv (a_N - 1)^2 + \frac{a_N^2}{\tau},$$
$$\sigma_{C,N}^2 \equiv (a_N - 1)^2 + \frac{a_N^2}{N\tau}.$$

Then the four information-quality measures are

$$\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^N |X_i - Z|\right] = 2\sigma_{I,N}\phi\left(\frac{b_N}{\sigma_{I,N}}\right) + b_N\left[2\Phi\left(\frac{b_N}{\sigma_{I,N}}\right) - 1\right], \tag{32}$$

$$\mathbb{E}\left[\frac{1}{N-1}\sum_{i=1}^N (X_i - X^*)^2\right] = \frac{a_N^2}{\tau}, \tag{33}$$

$$\mathbb{E}|X^* - Z| = 2\sigma_{C,N}\phi\left(\frac{b_N}{\sigma_{C,N}}\right) + b_N\left[2\Phi\left(\frac{b_N}{\sigma_{C,N}}\right) - 1\right], \tag{34}$$

$$\mathrm{Var}(X^*) = a_N^2\left(1 + \frac{1}{N\tau}\right). \tag{35}$$

Proof. Let $\bar{\epsilon} \equiv N^{-1}\sum_{i=1}^N \epsilon_i$. Substituting the signal structure into the individual and consensus forecasts gives

$$X_i - Z = (a_N - 1)Z + a_N\epsilon_i + b_N,$$
$$X^* - Z = (a_N - 1)Z + a_N\bar{\epsilon} + b_N.$$

Because Z and the signal errors are independent normal random variables and $\bar{\epsilon} \sim N(0, 1/(N\tau))$, the two forecast errors have distributions

$$X_i - Z \sim N\left(b_N, \sigma^2_{I,N}\right),$$
$$X^* - Z \sim N\left(b_N, \sigma^2_{C,N}\right).$$

For any $Y \sim N(\mu, \sigma^2)$, write $Y = \mu + \sigma U$ with $U \sim N(0,1)$. Splitting the expectation at $U = -\mu/\sigma$ gives

$$\begin{aligned}\mathbb{E}|Y| &= \int_{-\infty}^{-\mu/\sigma} -(\mu+\sigma u)\phi(u)\,du + \int_{-\mu/\sigma}^{\infty} (\mu+\sigma u)\phi(u)\,du \\ &= 2\sigma\phi\left(\frac{\mu}{\sigma}\right) + \mu\left[2\Phi\left(\frac{\mu}{\sigma}\right) - 1\right].\end{aligned}$$

Applying this identity to the two error distributions proves equations (32) and (34). Symmetry across analysts implies that the expected average individual forecast error equals the expected forecast error for any individual analyst.

Next, consider individual forecast dispersion. Since $X^* = a_N Z + a_N \overline{\epsilon} + b_N$, the common component $a_N Z + b_N$ cancels when individual forecasts are demeaned:

$$X_i - X^* = a_N(\epsilon_i - \overline{\epsilon}).$$

It follows that

$$\begin{aligned}\mathbb{E}\left[\frac{1}{N-1}\sum_{i=1}^{N}(X_i - X^*)^2\right] &= \frac{a_N^2}{N-1}\mathbb{E}\left[\sum_{i=1}^{N}(\epsilon_i - \overline{\epsilon})^2\right] \\ &= \frac{a_N^2}{N-1}\frac{N-1}{\tau} \\ &= \frac{a_N^2}{\tau}.\end{aligned}$$

The denominator $N-1$ makes this an unbiased measure of the idiosyncratic forecast variance for every finite N, proving equation (33).

Finally, the variance of the consensus forecast across contests is

$$\begin{aligned}\mathrm{Var}(X^*) &= a_N^2\,\mathrm{Var}(Z + \overline{\epsilon}) \\ &= a_N^2\left(1 + \frac{1}{N\tau}\right),\end{aligned}$$

which proves equation (35). ■

Conditional on the equilibrium coefficients (a_N, b_N), all four expressions are exact. The number of analysts affects the individual measures through its effect on the equilibrium strategy. It additionally affects the consensus measures directly through the reduction in signal noise from $1/\tau$ to $1/(N\tau)$. Equation (35) is the variance of consensus forecast levels across contests, not the variance of the consensus forecast error.

D Computation Details

D.1 Equilibrium Computation

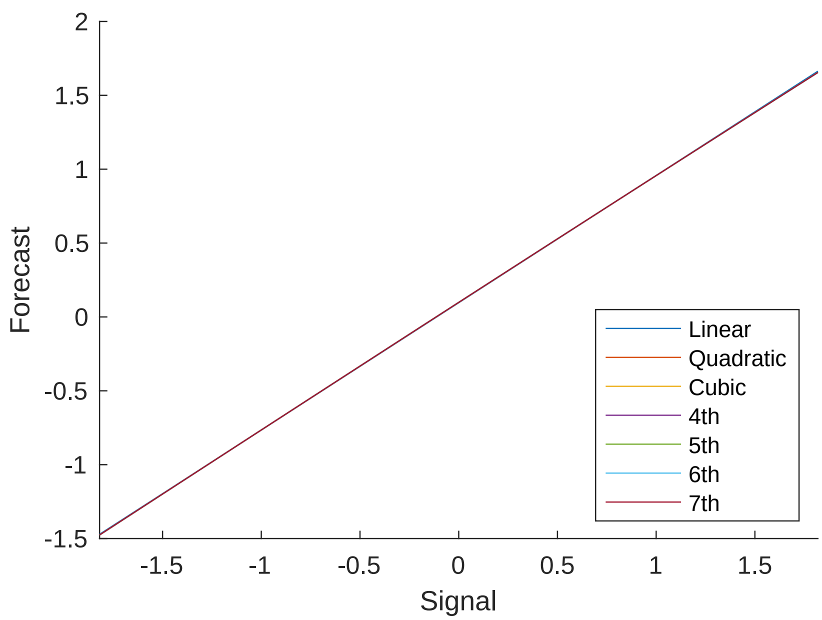
In this section, I provide details on equilibrium computation.

Experimentation with polynomial orders I experiment with different polynomial orders as the pre-specified function family in the projection method and show that a linear strategy is an excellent approximation to the equilibrium strategy.

I choose the representative signals (collocation points) to be the 20 evenly spaced interior quantiles of the signal distribution. Then, under the parametric assumption that $S_i = Z + \epsilon_i$ where $Z \sim N(0, 1)$ and $\epsilon_i \sim N(0, 1/\tau)$, these 20 collocation points are

$$\left\{ \sqrt{1 + \frac{1}{\tau}} \Phi^{-1}(\frac{1}{21}), \sqrt{1 + \frac{1}{\tau}} \Phi^{-1}(\frac{2}{21}), \cdots, \sqrt{1 + \frac{1}{\tau}} \Phi^{-1}(\frac{20}{21}) \right\}.$$

Figure D.1: Equilibrium Strategy Approximated With Polynomials up to the 7th Order



Note: Figure plots the equilibrium strategy approximated with polynomials from the 1st order to the 7th order.

The equilibria are computed following the algorithm described in Section 4.3. I experiment with polynomials up to the 7th order as the pre-specified function at various model

Table D.1: Equilibrium Approximation Performance of Polynomials up to the 7th Order

Polynomial Order	Residual Function	Computation Time (s)
1	2.21e-02	0.26
2	4.07e-04	0.86
3	2.39e-04	1.68
4	2.37e-04	2.49
5	9.97e-10	4.64
6	9.49e-10	6.07
7	9.49e-10	9.03

Note: Table presents performance of polynomials from the 1st order to the 7th order in approximating the equilibrium strategy, measured by residual function and computation time in seconds.

parameters and compare the performance of these different polynomial orders using the residual function, i.e., the sum of squared first-order conditions.

Figure D.1 plots the approximated equilibria with 10 players using polynomials up to the 7th order at the estimated parameters in Table 3. Table D.1 presents their corresponding residual functions and computation time with MATLAB on a macOS machine with 16 GB of RAM and an Apple M1 chip. Higher polynomial orders generally achieve smaller residual functions, but the estimated polynomials are very close to the linear function and cost much more computation time. This is observed consistently at different model parameters. Therefore, I use a linear function to approximate the equilibrium in both estimation and counterfactual analysis.

Interpolation of equilibrium strategy I interpolate the equilibrium forecast strategy to alleviate the computation burden in the structural estimation. For a given set of parameter values, I first use the projection method to compute the equilibria at every 5th percentile of the number of analysts N. Then, for each coefficient in the equilibrium strategy, I use a cubic spline to interpolate across N to find the equilibrium strategy with any other number of analysts.

D.2 Data Homogenization

This section details how I homogenize the data for the structural estimation. I regress analysts' forecasts X_{ijt}, the number of analysts N_{jt}, and the realized earnings Z_{jt} on security fixed effects and year fixed effects, and take the residuals as the homogenized variables, denoted by the h superscript. This procedure implicitly assumes that the unobserved heterogeneity in these variables is additively separable and lies only at the security level and the year level. That is,

$$
\begin{aligned}
X_{ijt} =& X_{ijt}^h + \delta_j^X + \delta_t^X \\
N_{jt} =& N_{jt}^h + \delta_j^N + \delta_t^N \\
Z_{jt} =& Z_{jt}^h + \delta_j^Z + \delta_t^Z,
\end{aligned}
$$

where the δ_j's and δ_t's are unobserved heterogeneity at the security level and the year level, respectively. Removing them allows me to treat the homogenized variables as outcomes of the homogeneous contests described in Section 4.

The homogenized variables still vary within a security over time, and part of this variation may remain endogenous. To fix ideas, consider the following Markov process for the homogenized number of analysts:

$$
N_{jt}^h = \rho_N(N_{jt-1}^h, \delta_{jt}^N),
$$

where δ_{jt}^N is the shock on the number of analysts covering security j in year t. The homogenized forecast X_{ijt}^h is a function of the analyst's signal and may be subject to time-varying unobserved heterogeneity δ_{jt}^X such that

$$
X_{ijt}^h = \rho_X(S_{ijt}, \delta_{jt}^X).
$$

If δ_{jt}^N and δ_{jt}^X are correlated, the homogenized number of analysts will be endogenous in equations (9) to (11), so I instrument it with its one-year lag N_{jt-1}^h. This implicitly assumes that the lagged number of analysts is uncorrelated with the current-period shocks δ_{jt}^N and δ_{jt}^X.

After the homogenization, I keep the security-years covered by at least two homogenized analysts, which accounts for 99.6% of the sample, so that the contest remains well-defined. With a slight abuse of notation, I drop the h superscript and use $(X_{ijt}, Z_{jt}, N_{jt})$ to denote the homogenized variables from Section 5.2.1 onward in the paper.

E Properties of the Specification With Flexible Convexity

In this section, I illustrate properties of the alternative specification where I allow flexible convexity in the relative accuracy reward. I set

$$
\gamma_k(N) \equiv v(k, N) = \gamma_1(1 - \frac{k-1}{N-1})^{\frac{1}{\gamma_{convex}}},
$$

where $\gamma_1 \geq 0$ and $\gamma_{convex} > 0$. Note that $\gamma_1(N) = \gamma_1$, so γ_1 is the payoff for the most accurate. Rewards at other ranks are scaled by $(1 - \frac{k-1}{N-1})^{\frac{1}{\gamma_{convex}}}$, which is decreasing in k.

Proposition 4. Suppose $\gamma_1 > 0$ and $N \geq 3$. Consider any integers k and l such that $1 \leq k < l \leq N-1$. (a) If $\gamma_{convex} \in (0,1)$, $\gamma_k(N) - \gamma_{k+1}(N) > \gamma_l(N) - \gamma_{l+1}(N)$; if $\gamma_{convex} \in (1, \infty)$, $\gamma_k(N) - \gamma_{k+1}(N) < \gamma_l(N) - \gamma_{l+1}(N)$; if $\gamma_{convex} = 1$, $\gamma_k(N) - \gamma_{k+1}(N) = \gamma_l(N) - \gamma_{l+1}(N)$. (b) For every interior rank $k \in \{2, \ldots, N-1\}$, $\gamma_k(N)$ is strictly increasing in γ_{convex}, while $\gamma_1(N) = \gamma_1$ and $\gamma_N(N) = 0$ for all $\gamma_{convex} > 0$.

Proof. (a) The function $v(\cdot, N)$ is continuous on $[k, k+1]$ and $[l, l+1]$ and differentiable on the interior of each interval. Therefore, by the mean value theorem, there exist $k^* \in (k, k+1)$ and $l^* \in (l, l+1)$ such that

$$\begin{aligned} v(k, N) - v(k+1, N) &= -v'(k^*, N), \\ v(l, N) - v(l+1, N) &= -v'(l^*, N). \end{aligned}$$

Because k and l are integers and $k < l$, $k^* < k+1 \leq l < l^*$.

Differentiating $v(k, N)$ twice with respect to k, we have

$$\frac{\partial^2 v(k,N)}{\partial k^2} = \frac{\gamma_1}{\gamma_{convex}(N-1)^2}\left(\frac{1}{\gamma_{convex}} - 1\right)\left(1 - \frac{k-1}{N-1}\right)^{\frac{1}{\gamma_{convex}} - 2}.$$

Note that when $\gamma_{convex} \in (0,1)$, $\frac{\partial^2 v(k,N)}{\partial k^2} > 0$, so v' is increasing. As $l^* > k^*$,

$$\begin{aligned} & v'(l^*, N) > v'(k^*, N) \\ \Leftrightarrow & -v'(k^*, N) > -v'(l^*, N) \\ \Leftrightarrow & v(k, N) - v(k+1, N) > v(l, N) - v(l+1, N) \\ \Leftrightarrow & \gamma_k(N) - \gamma_{k+1}(N) > \gamma_l(N) - \gamma_{l+1}(N). \end{aligned}$$

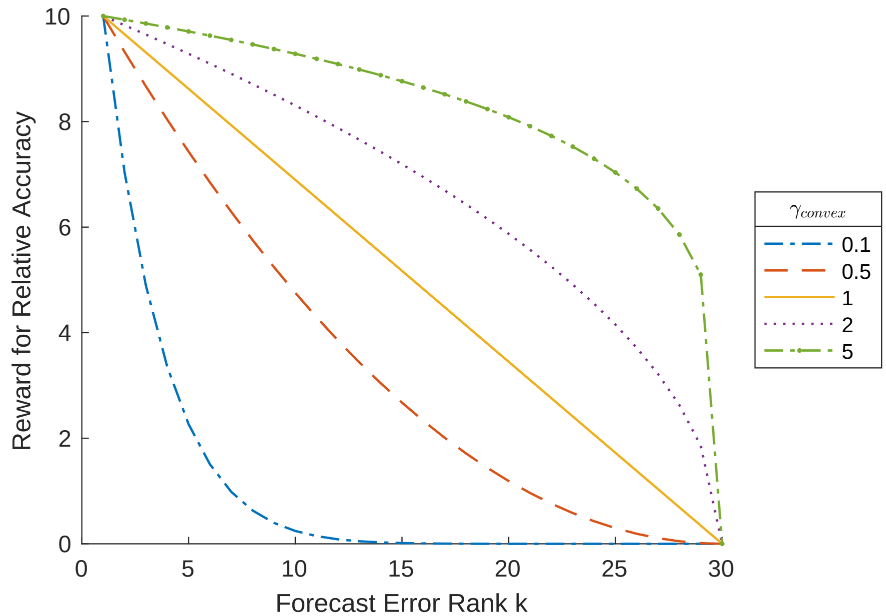
Similarly, when $\gamma_{convex} \in (1, \infty)$, $\frac{\partial^2 v(k,N)}{\partial k^2} < 0$, so v' is decreasing and $\gamma_k(N) - \gamma_{k+1}(N) < \gamma_l(N) - \gamma_{l+1}(N)$; when $\gamma_{convex} = 1$, $\frac{\partial^2 v(k,N)}{\partial k^2} = 0$, so v' is constant and $\gamma_k(N) - \gamma_{k+1}(N) = \gamma_l(N) - \gamma_{l+1}(N)$.

(b) For every $k \in \{2, \ldots, N-1\}$, let $a_k = 1 - (k-1)/(N-1) \in (0,1)$. Then,

$$\frac{\partial \gamma_k(N)}{\partial \gamma_{convex}} = -\frac{\gamma_1}{\gamma_{convex}^2} a_k^{1/\gamma_{convex}} \log(a_k) > 0.$$

The inequality follows because $\gamma_1 > 0$, $\gamma_{convex} > 0$, and $\log(a_k) < 0$. The endpoint rewards follow directly from the payoff specification. ■

Figure E.1: Relative Accuracy Payoffs with Different Convexity Parameters When $N = 30$



Note: Figure plots the relative accuracy payoffs with different convexity parameters when there are 30 analysts in the market. The reward for the most accurate is set at 10.

Proposition 4 shows how γ_{convex} determines the shape of the payoff function. The reward schedule is convex in rank when $\gamma_{convex} < 1$, linear when $\gamma_{convex} = 1$, and concave when $\gamma_{convex} > 1$. With the payoffs fixed for the most and the least accurate, lower values of γ_{convex} imply a more top-heavy relative accuracy payoff. To illustrate, Figure E.1 plots the relative accuracy payoffs at different values of γ_{convex} when there are 30 players. As $\gamma_{convex} \to 0$, this payoff approaches the winner-takes-all specification. As $\gamma_{convex} \to \infty$, this approaches the loser-loses-all specification.