Revista Colombiana de Estadística - Applied Statistics
January 2025, volume 48, issue 1, pp. 157–175
http://doi.org/10.15446/rce.v48n1.112678

Forecast Quality with Information Uncertainty. Balancing and Estimation from Propensity Scores

Predicciones con incertidumbre informacional. Balance y estimación a partir de puntajes de propensión

José Gabriel Astaíza-Gómez[a]

Área de Macroeconomía y Sistemas Financieros, Escuela FEG, Universidad EAFIT, Medellín, Colombia

Abstract

This study investigates whether analysts provide informative forecasts for stocks issued by firms with greater information uncertainty. As firm-specific information uncertainty is not directly observable, the research highlights the role of analysts' forecasts and reports in offering valuable insights to investors. It also investigates whether forecast quality is sufficiently captured by forecast bias. The findings indicate that forecast quality tends to be lower for firms with greater information uncertainty and that forecast bias alone does not fully reflect the informational content of analysts' forecasts. Overall, the results suggest that analysts' forecasts possess positive informational value.

***Key words*:** Error consistency; Propensity scores; Risk; Value ambiguity.

Resumen

Este estudio investiga si los analistas proporcionan pronósticos informativos para las acciones emitidas por empresas con mayor incertidumbre de información. Dado que la incertidumbre de información específica de la empresa no es directamente observable, el artículo destaca el papel de los pronósticos y los informes de los analistas en ofrecer información valiosa a los inversores. También examina si la calidad de los pronósticos se captura adecuadamente mediante el sesgo de los pronósticos. Los resultados indican que la calidad de los pronósticos tiende a ser menor para las empresas con mayor incertidumbre de información y que el sesgo de los pronósticos por sí solo no refleja completamente el contenido informativo de estos. En conjunto, los resultados sugieren que los pronósticos de los analistas poseen un valor informativo positivo.

***Palabras clave*:** Ambigüedad de valor; Consistencia del error; Puntuaciones de propensión; Riesgo.

[a]Ph.D. E-mail: jastaiza@eafit.edu.co

1. Introduction

Risk is a primary consideration in portfolio management. Traditional asset pricing literature assumes that the existing informational environment provides adequate inputs for a risk-return framework. In this framework, ambiguity regarding the estimation of a firm's fundamental value—whether due to the volatility of the firm's fundamentals or the presence of hard-to-observe risks—is not typically addressed. This ambiguity, known as information uncertainty, is crucial for investment decisions (Kang et al., 2019; Jiang et al., 2005; Zhang, 2006) and differs from investment risk. Greater information uncertainty complicates investors' ability to make informed decisions about the associated risks (Hansen & Sargent, 2010; Morgan, 2002).

The market's response to increased information uncertainty is reflected in lower future stock returns (Jiang et al., 2005) and greater investment mistakes (Kumar, 2009). Analysts, as information intermediaries, offer stock valuations that may provide insights into firms' information uncertainty. Specifically, sell-side analysts gather and interpret costly information, which they use to issue stock valuations that assist investors in their decision-making (Huang et al., 2014; Hilary & Hsu, 2013). However, given that analysts have commercial incentives (Cowen et al., 2006; Jackson, 2005) and often issue optimistic forecasts (Malmendier & Shanthikumar, 2014; Lim, 2001), it remains uncertain whether their valuations effectively illuminate a firm's information uncertainty.

I investigate the extent to which the dynamics of analyst valuations reflect information uncertainty, and how the tendency to issue optimistic forecasts influences the effectiveness of analyst valuations in guiding investment decisions. Analysts forecasts are informative as long as investors can unravel consistent errors, which correspond to systematic biases, but if analysts forecasts are not a predictable transformation of realized values, these are empty of informational content. That is, forecast quality is higher to the degree in which forecasts are a predictable transformation of realized values and the higher the standard deviation of forecast bias, the lower the informativeness of forecasts (Hilary & Hsu, 2013). I find that forecast quality is statistically lower for firms of greater information uncertainty. As investors have varying risk preferences, a variable capturing information uncertainty allows portfolio managers to better tailor investment strategies to align with the risk preferences and objectives of their clients. I also find that forecast bias is lower for these firms. Uncertainty makes an additional contribution to the price of risk (Hansen & Sargent, 2010) and the information uncertainty of opaque assets entail higher valuation discounts relative to transparent assets (Jones et al., 2013).

My first contribution is to analyze the informational content of analysts' forecasts on firm information uncertainty. This is important because information uncertainty is a relevant factor motivating a wide variety of recent financial outcomes[1]. Firm information uncertainty is not an observable variable and empirical

[1]For instance, on lottery-like stocks (Tao et al., 2020), investor underreaction (Jia et al., 2020) and enterprise systems portfolios (Sambhara et al., 2022)

researchers use proxies based on a measurable firm characteristic to test their hypothesis. Analysts' forecasts and reports are readily available for investors and there are direct advantages of understanding their usefulness for capturing information uncertainty around stock expectations. Second, I addresses the role of analysts as information intermediaries in the stock market. My findings align with literature indicating that information gathering in the stock market is costly and provide evidence supporting the informational value of analysts' forecasts. Despite being, on average, positively biased, these forecasts are valuable for which analysts are compensated. The extent to which analyst coverage alone is a good proxy for the informational environment of firms is an interesting subject of further research. My results offer an explanation for the mixed empirical results regarding the relationship between analyst coverage and firm opacity.

This paper contains five sections including the introduction. In section two I describe the literature related to firm information uncertainty and sell-side analysts. Next, in sections three and four, I describe the methodology and present the results. Finally in section five, I conclude.

2. Related Literature

Incorporating information uncertainty is important for asset allocation as it is significantly correlated with illiquidity (Kang et al., 2019) and lower future stock returns (Jiang et al., 2005). More recently, Hao et al. (2024) measure information uncertainty using satellite-based estimates of oil inventory and find that higher information uncertainty is associated with lower future stock returns. Additionally, Benamar et al. (2021) demonstrate that the number of clicks on short-URL links correlates with the implied and realized volatility of Treasury note returns.

Analysts' valuations help investors to form expectations for projected returns (Cheng et al., 2006) but potentially these forecasts have an additional piece of information related to information uncertainty which complements the attempts to take decisions in the risk-return space. This is an opportunity to improve portfolio diversification as estimating the inputs for optimization based on market historical information is not an easy task (Maccheroni et al., 2013; DeMiguel & Nogales, 2009). As outsiders, analysts may not be capable of issuing forecasts with informational content on firms of greater information uncertainty. Fischer & Stocken (2010) theoretically find that, if the costs of gathering and interpreting information are too high, analysts give up in their tasks, thus providing forecasts without informational content or providing no forecasts at all[2]. The difficulty of evaluating the prospects of a firm and the value of its stock is related to the historical variability of its earnings, which is a variable related to fundamental value, and to the volatility of its stock returns (Aslan & Kumar, 2017). Also, stock valuation is dependent upon the amount of information about a company that is publicly available (Joos et al., 2016; Bilinski et al., 2013; Lim, 2001; Wieland, 2011; Lys & Soo, 1995).

[2]See Proposition 5.

Forecast error consistency provides a good framework to study whether analysts' reports properly reflect the degree of information uncertainty of firms. Consistency is measured by the standard deviation of forecast bias, and is in line with the intuition that investors are able to debias forecasts that are systematically biased, notwithstanding the size of the bias. This is supported by the empirical evidence in Hilary & Hsu (2013): more consistent forecast errors move stock prices to a greater extent and this effect increases with the presence of sophisticated investors. What is more, the effect of consistency is two to four times larger than the effect of accuracy. Forecast bias in turn, is unlikely to account for the level of informational content on analysts forecasts because analysts are able to report their beliefs untruthfully (Fischer & Stocken, 2010; Beyer & Guttman, 2011), driven mainly by their trading (Cowen et al., 2006; Jackson, 2005) and reputational incentives (Groysberg et al., 2011; Mikhail et al., 1999).

Forecasts on stock prices or target prices can circumvent the issues associated with earnings management that arise when using earnings forecasts to detect forecast bias or errors. The quality of financial statements may be affected by managerial actions, such as altering the estimation methods for accruals and other accounting metrics (Beaver, 2002). Inappropriately benchmarking forecast errors against manipulated earnings introduces errors unrelated to analysts' skills (Abarbanell & Lehavy, 2003). Additionally, the market tends to react more strongly to revisions in target prices compared to changes in earnings forecasts (Asquith et al., 2005).

3. Methodology

3.1. Data and Variables

For 2,695 firms listed in the Center for Research in Security Prices (CRSP) stock index from Q2 2006 to Q4 2017, I analyze quarterly Earnings Per Share (EPS) data, along with stock price and market capitalization. Additionally, I examine the consensus target price, which is the average forecast of the stock price for the next 12 months from analysts covering the stock, excluding forecasts older than three months. My dependent variables are forecast bias ($\bar{y}_i$) and forecast quality (γ_i). For each firm i, I calculate forecast bias as the logarithm of the time-mean of

$$y_{i,t} = \frac{TP_{i,t-4} - P_{i,t}}{P_{i,t-4}}$$

where $TP_{i,t-4}$ is the consensus forecast or target price, issued at the end of the quarter $t-4$ for the next 4 quarters on stock i and $P_{i,t}$ is the stock price at the end of the quarter t. To better assess analysts' ability to predict future stock price movements, I use the covariance between projected and realized returns, termed forecast covariance, instead of the traditional standard deviation. The standard deviation or variance of forecast bias captures not only analysts' prediction ability (through covariance) but also the volatility of returns due to market conditions unrelated to analysts' efforts. For each firm i, the time-variance of forecast bias (consistency) is

$$\begin{aligned} Var_i[y_{it}] &= Var_i\left[\frac{TP_{i,t-4}}{P_{i,t-4}} - \frac{P_{i,t}}{P_{i,t-4}}\right] \\ &= Var_i\left[r^p_{i,t-4} - r_{i,t}\right] \end{aligned} \tag{1}$$

where $r^p_{i,t-4}$ are analysts' projected returns and $r_{i,t}$ are realized returns. Since $\text{Var}_i\left[r^p_{i,t-4} - r_{i,t}\right] = \text{Var}_i\left[r^p_{i,t-4}\right] + \text{Var}_i\left[r_{i,t}\right] - 2\text{Cov}\left[r^p_{i,t-4}, r_{i,t}\right]$, a higher covariance indicates a lower variance of forecast bias, which suggests higher forecast quality. I calculate the logarithm of the covariance between the projected returns $r^p_{i,t-4}$ and the realized returns $r_{i,t}$ for each stock i, denoted as γ_i.

Financial institutions have greater information uncertainty as their business lines, fundings sources, high leverage and the nature of their assets, entail financial information that is not easy to gather or interpret. The limited assets that are physically fixed, opaque borrowers and the level of loan diversification of the lending firms, create collateral uncertainty that is hard to assess (Morgan, 2002) since each loan cannot be individually examined by an analyst. Moreover, the relative importance of their business lines such as investment banking or commercial banking (Banerji & Basu, 2017), and of their funding sources such as retail deposits or wholesale financiers (Calomiris & Kahn, 1991; Huang & Ratnovski, 2011), as well as the differences in deposit insurance schemes around the world (Matutes & Vives, 1996), entail risks for these firms that are not easy to understand for outsiders, making financial institutions more opaque. Correspondingly, the empirical literature finds that there is greater heterogeneity in bond ratings for financial institutions than for non-financials (Morgan, 2002) and that stocks issued by banks exhibit more inefficient prices (Blau et al., 2017; Dahiya et al., 2017). I differentiate opaque firms with higher information uncertainty by identifying those stocks issued by financials. I calculate the dummy D_i^{Fin} that takes the value of one whenever the stock i belongs to the financial sector, as defined in the CRSP Financials Index (NASDAQ symbols CRSPFN1 and CRSPFNT).

Technology firms count with higher amounts of firm information provided by a wide range of media sources (Bartov et al., 2018; Corea, 2016; Greenwood & Gopal, 2015) and sell-side analysts are better at assessing the risks of technology companies (Forbes et al., 2020; Su et al., 2019). I calculate D_i^{Tech} which takes the value of one whenever the stock belongs the the technology sector, as defined in the CRSP Technology Index. There are 10 CRSP sector indexes, including Technology and Financials. The Financials index consists of 832 firms, including banks, credit card companies, trading companies, and real estate investment trusts (REITs). The Technology index includes 403 firms, such as mobile producers, chip manufacturers, software developers, and manufacturers of communications and measuring equipment.

Among the variables explaining forecast quality, I include earnings variability, estimated as the log[3] of the standard deviation of changes in Earnings Per Share (EPS) scaled by the stock price, i.e., the log of the standard deviation of $\frac{EPS_{i,t}-EPS_{i,t-1}}{P_{i,t-1}}$, and the log of the standard deviation of quarterly stock returns.

[3]The derivative of the logarithm of a variable equals the percentage change of the variable, and the dependent variable is a logarithm. Therefore, the coefficients on the log of volatilities account for elasticities.

I denote these variables as σ_i^{eps} and σ_i^{retn}, respectively. Following the literature, I also include firm size ($size_i$) as a proxy for the informational environment of firms. Summary statistics are provided in Table 1. As shown, approximately 21% of the firms in the dataset belong to the financial sector, and approximately 11% belong to the technology sector.

TABLE 1: Summary statistics.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
γ_i	-1.5815	-0.0001	0.0080	0.0396	0.0308	3.1561
$\bar{y}_i$	-2.3768	-0.0271	0.0598	0.1452	0.2424	1.99812
σ_i^{eps}	-7.140	-4.584	-3.663	-3.510	-2.522	4.731
σ_i^{retn}	-4.0375	-1.3687	-1.0087	-0.9559	-0.5699	2.5336
$size_i$	2.428	5.811	7.006	7.097	8.283	12.863
D_i^{Fin}	0.00	0.00	0.00	0.22	0.00	1.00
D_i^{Tech}	0.0000	0.0000	0.0000	0.1076	0.0000	1.0000

For each firm i, γ_i is the log of the covariance between projected and realized returns on stock i; $\bar{y}_i$ is the log of the time-average of forecast bias; D_i^{Fin} is a dummy that takes the value of one whenever the stock i belongs the CRSP Financials Index; D_i^{Tech} takes the value of one whenever the stock i belongs the CRSP Technology Index; σ_i^{eps} is the log of the standard deviation of changes in Earnings Per Share; σ_i^{retn} is the log of the standard deviation of stock returns; $size_i$ corresponds to the log of the time-mean of market capitalization.

3.2. Propensity Scores and Estimation

I estimate the following linear model:

$$\gamma_i = \alpha + \tau D_i^{Fin} + \boldsymbol{x}_i\boldsymbol{\theta}_0 + \boldsymbol{x}_i D_i^{Fin}\boldsymbol{\theta}_1 + \varepsilon_i \tag{2}$$

where γ_i is the forecast quality on stock i. The vector $\boldsymbol{x}_i$ of 1×3 contains the explanatory variables σ_i^{eps} and σ_i^{retn}, as well as $size_i$. Additionally, $\boldsymbol{\theta}_0$ is the vector of parameters on regressors, $\boldsymbol{\theta}_1$ is the vector of parameters on interactions and ε_i is the error term. In this equation, the parameter τ provides the estimated change of forecast quality conditional to variations on information uncertainty.

I group treated and control units so that direct comparisons are more meaningful, using a balancing score. If $\gamma_{i,0}$ is the covariance between projected and realized returns for a non-financial firm of certain characteristics and $\gamma_{i,1}$ is the covariance for a financial firm, for the same firm I only observe $\gamma_{i,0}$ when $D_i^{Fin} = 0$ and $\gamma_{i,1}$ when $D_i^{Fin} = 1$. However, I do not observe the covariance of treated firms, had they not been treated, i.e. $\gamma_{i,1}$ when $D_i^{Fin} = 0$. What I observe is (Quandt, 1958, 1972)

$$\gamma_i = D_i^{Fin}\gamma_{i,1} + (1 - D_i^{Fin})\gamma_{i,0} \tag{3}$$

I run OLS regressions using matched data from propensity scores. As Rosenbaum and Rubin (1983) showed, given the propensity score, the treatment is exogenous. I use propensity scores of logit regressions and match treated units to control units that are closest in terms of the distance of their scores, in order to obtain similar conditional distributions of $\boldsymbol{x}_i$ for treated and untreated units.

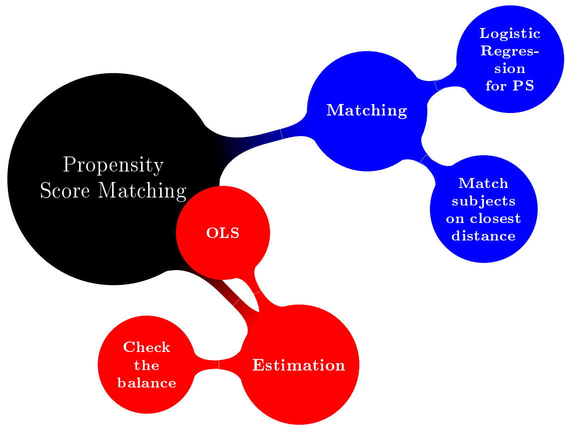

Figure 1: Propensity Score Matching.

4. Results

In Table 2 I show the results on different specifications using γ_i as the dependent variable. The estimates are negative and statistically significant on the variable for information uncertainty. Analysts' forecasts on firms with higher information uncertainty, follow the realized values of stock prices to a lesser extent for opaque firms of higher information uncertainty.

Table 2: OLS estimates of the effect of financial firms on forecast covariance.

	(1)	(2)	(3)	(4)
D_i^{Fin}	-0.035^{***}	-0.098^{***}	-0.097^{***}	-0.121^{***}
	(0.008)	(0.023)	(0.024)	(0.038)
σ_i^{eps}		0.025^{***}	0.006^{**}	0.005^{*}
		(0.003)	(0.003)	(0.003)
σ_i^{retn}			0.092^{***}	0.089^{***}
			(0.006)	(0.007)
$size_i$				-0.003
				(0.002)
Constant	0.047^{***}	0.132^{***}	0.148^{***}	0.161^{***}
	(0.004)	(0.010)	(0.009)	(0.015)
Observations	2,695	2,695	2,695	2,695

Forecast covariance (γ_i) is my dependent variable and opacity (D_i^{Fin}) is my treatment variable. For each firm i, γ_i is the log of the covariance between projected and realized returns on stock i; D_i^{Fin} is a dummy that takes the value of one whenever the stock i belongs the CRSP Financials Index; σ_i^{eps} is the log of the standard deviation of changes in Earnings Per Share; σ_i^{retn} is the log of the standard deviation of stock returns. Variables related to the informational environment are $size_i$, which corresponds to the log of the time-mean of market capitalization; $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

These results are interesting as previous empirical results show that there is more to the relationship between analyst's reports and firm opacity than forecast accuracy or the plain amount of information issued on firms with information uncertainty. Flannery et al. (2004) argue that lower forecast errors is an indication of higher forecast quality and Derrien & Kecskes (2013) and Mehran & Peristiani (2009), assume that higher analyst coverage implies higher firm transparency. By contrast, other research find a positive relation between the level of analyst coverage and firm opacity (Crawford et al., 2012; Chan & Hameed, 2006; Piotroski & Roulstone, 2004). To complete the story, I run OLS regressions using matched data and forecast bias, $\bar{y}_i$ as my dependent variable.

The propensity score matching balances the covariates between the firms in the Financials index and firms outside the index. I show the results of this procedure in Table 3. There is an improvement in balance for all variables resulting in similar means for treated and untreated units. Also, Figure 2 shows that the distribution of the propensity scores becomes similar after the matching. I estimate the mean of matched pair differences, which is an unbiased estimator of the average treatment effect[4]. I estimate the mean differences through OLS regressions using the matched observations with the same specifications as before. I show the estimates in Table 4.

TABLE 3: Summary of balance for financial firms.

	All data		Matched data		Balance improvement
	Means Treated	Means Control	Means Treated	Means Control	
σ_i^{eps}	-3.8528	-3.4137	-3.8528	-3.7863	84.8647%
σ_i^{retn}	-1.2730	-0.8664	-1.2730	-1.2014	82.3869%
$size_i$	7.1762	7.0751	7.1762	7.2496	27.4177%

From the 2,695 stocks in the sample, 593 correspond to the Financials index. In the matched sample, both the control and treatment groups are composed of 593 units. The number of unmatched observations is 1,509. *Improvement* corresponds to the percentage change of the difference of averages between the entire sample and the matched data. The final matched sample does not include $volm_i$ since it was not possible to improve the balance using this variable.

[4]See Corollary 4.1 of Rosenbaum & Rubin (1983).

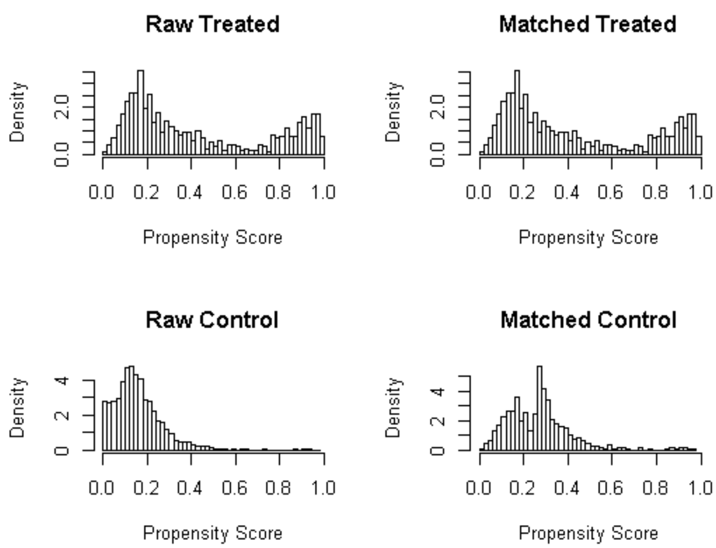

Figure 2: Histograms of propensity scores for financial firms.

All estimates on D_i^{Fin} are statistically significant and negative, indicating that analysts issue forecasts with less informational content for opaque firms. Also, in all specifications, the variability on Earnings Per Share and stock returns are statistically significant and negatively associated to γ_i for firms in the financial sector.

I investigate further into the relationship between analysts' reports and information uncertainty using forecast errors. The results are in Table 5. Analysts are less optimistic for firms of greater information uncertainty (negative estimates on D_i^{Fin}) indicating that forecast quality provides a different piece of information than forecast error. I contrast the results on firms with higher information uncertainty, regressing forecast quality on D_i^{Tech} as there is more information on technology firms provided by free sources and empirical analyses show that sell-side analysts are better at assessing their risks. I use matched data for the regressions. In Table 6 and Figure 3 I show the results of a propensity score matching procedure using those firms in the Technology index as the treated firms.

Table 4: OLS results with matched data for financial firms.

	(1)	(2)	(3)	(4)
D_i^{Fin}	-0.010	-0.055^{***}	-0.062^{***}	-0.066^{**}
	(0.006)	(0.018)	(0.018)	(0.026)
σ_i^{eps}		0.018^{***}	0.007^{*}	0.007^{**}
		(0.003)	(0.003)	(0.004)
σ_i^{retn}			0.054^{***}	0.056^{***}
			(0.008)	(0.008)
$size_i$				0.002
				(0.002)
$D_i^{Fin} * \sigma_i^{eps}$		-0.012^{***}	-0.006	-0.007
		(0.004)	(0.005)	(0.005)
$D_i^{Fin} * \sigma_i^{retn}$			-0.025^{**}	-0.027^{**}
			(0.011)	(0.012)
$D_i^{Fin} * size_i$				0.00001
				(0.003)
Constant	0.022^{***}	0.089^{***}	0.113^{***}	0.105^{***}
	(0.004)	(0.013)	(0.013)	(0.016)
Observations	1,186	1,186	1,186	1,186

Forecast covariance (γ_i) is my dependent variable. For each firm i, γ_i is the log of the covariance between projected and realized returns on stock i; D_i^{Fin} is a dummy that takes the value of one whenever the stock i belongs the CRSP Financials Index; σ_i^{eps} is the log of the standard deviation of changes in Earnings Per Share; σ_i^{retn} is the log of the standard deviation of stock returns. Variables related to the informational environment of firms are $size_i$, which corresponds to the log of the time-mean of market capitalization. From the 2,695 stocks in the sample, 593 correspond to the Financials index. In the matched sample, both the control and treatment groups are composed of 593 units. The number of unmatched observations is 1,509. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

TABLE 5: OLS results with matched data for financial firms.

	(1)	(2)	(3)	(4)
D_i^{Fin}	-0.098^{***}	-0.288^{***}	-0.337^{***}	-0.529^{***}
	(0.014)	(0.036)	(0.038)	(0.052)
σ_i^{eps}		0.106^{***}	0.092^{***}	0.070^{***}
		(0.006)	(0.007)	(0.007)
σ_i^{retn}			0.064^{***}	0.010
			(0.016)	(0.015)
$size_i$				-0.053^{***}
				(0.004)
$D_i^{Fin} * \sigma_i^{eps}$		-0.051^{***}	-0.031^{***}	-0.010
		(0.009)	(0.010)	(0.010)
$D_i^{Fin} * \sigma_i^{retn}$			-0.102^{***}	-0.048^{**}
			(0.023)	(0.023)
$D_i^{Fin} * size_i$				0.046^{***}
				(0.007)
Constant	0.133^{***}	0.532^{***}	0.560^{***}	0.793^{***}
	(0.010)	(0.026)	(0.026)	(0.032)
Observations	1,186	1,186	1,186	1,186

Forecast bias ($\bar{y}_i$) is my dependent. For each firm i, $\bar{y}_i$ is the log of the time-mean of forecast bias; D_i^{Fin} is a dummy that takes the value of one whenever the stock i belongs the CRSP Financials Index; σ_i^{eps} is the log of the standard deviation of changes in Earnings Per Share; σ_i^{retn} is the log of the standard deviation of stock returns. Variables related to the informational environment of firms are $size_i$, which corresponds to the log of the time-mean of market capitalization. From the 2,695 stocks in the sample, 593 correspond to the Financials index. In the matched sample, both the control and treatment groups are composed of 593 units. The number of unmatched observations is 1,509. $^*p < 0.1$; $^{**} < 0.05$; $^{***}p < 0.01$.

Table 6: Summary of balance for technology firms.

	All data		Matched data		Balance improvement
	Means Treated	Means Control	Means Treated	Means Control	
σ_i^{eps}	-3.7196	-3.4851	-3.7196	-3.6681	78.0550%
σ_i^{retn}	-0.8281	-0.9713	-0.8281	-0.8106	87.8053%
$size_i$	6.9626	7.1136	6.9626	7.0279	56.7609%

From the 2,695 stocks in the sample, 290 correspond to the Technology index. In the matched sample, both the control and treatment groups are composed of 290 units. The number of unmatched observations is 2115. *Improvement* corresponds to the percentage change of the differences of averages between the entire sample and the matched data.

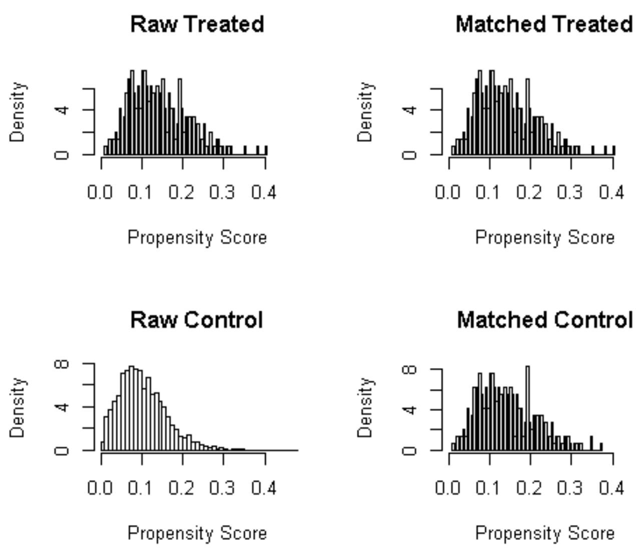

Figure 3: Histograms of propensity scores for technology firms.

With this matched data, I run OLS regressions using forecast bias, $\bar{y}_i$, as my dependent variable and D_i^{Tech} as the independent variable of interest. The results are in Table 7. While the estimates on D_i^{Fin} of Table 5 are negative, the positive estimates on D_i^{Tech} of Table 7 indicate that stock valuations are higher for firms when there is a greater amount of information circulating on the Internet and when analysts excel at valuing them.

Table 7: OLS results with matched data for technology firms.

	(1)	(2)	(3)	(4)
D_i^{Tech}	0.030	0.218***	0.224***	0.144
	(0.028)	(0.070)	(0.070)	(0.097)
σ_i^{eps}		0.064***	0.067***	0.043***
		(0.013)	(0.015)	(0.014)
σ_i^{retn}			−0.010	−0.065**
			(0.032)	(0.030)
$size_i$				−0.082***
				(0.011)
$D_i^{Tech} * \sigma_i^{eps}$		0.050***	0.043**	0.025
		(0.018)	(0.020)	(0.019)
$D_i^{Tech} * \sigma_i^{retn}$			0.040	0.018
			(0.047)	(0.044)
$D_i^{Tech} * size_i$				−0.001
				(0.015)
Constant	0.107***	0.343***	0.343***	0.787***
	(0.020)	(0.050)	(0.051)	(0.074)
Observations	580	580	580	580

Forecast bias ($\bar{y}_i$) is my dependent variable and firms in the Technology sector (D_i^{Tech}) is my treatment variable. For each firm i, $\bar{y}_i$ is the log of the time-mean of forecast bias; D_i^{Tech} is a dummy that takes the value of one whenever the stock i belongs the CRSP Technology Index; σ_i^{eps} is the log of the standard deviation of changes in Earnings Per Share; σ_i^{retn} is the log of the standard deviation of stock returns. Variables related to the informational environment of firms are $size_i$, which corresponds to the log of the time-mean of market capitalization. From the 2,695 stocks in the sample, 290 correspond to the Technology index. In the matched sample, both the control and treatment groups are composed of 290 units. The number of unmatched observations is 2115. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

In Table 8 I show the results of regressing forecast covariance on D_i^{Tech}. The idea that forecast bias does not account for the level of informational content on higher information uncertainty firms is plausible. Analysts are able to report their beliefs untruthfully, driven mainly by their trading and reputational incentives. The positive estimates on D_i^{Tech} show that analysts do issue forecasts with informational content on technology firms, notwithstanding their higher levels of forecast optimism on these same firms.

Table 8: OLS results with matched data for technology firms.

	(1)	(2)	(3)	(4)
D_i^{Tech}	-0.006	-0.001	0.012	0.099^{**}
	(0.012)	(0.034)	(0.031)	(0.047)
σ_i^{eps}		0.015^{**}	-0.015^{**}	-0.012^{*}
		(0.006)	(0.007)	(0.007)
σ_i^{retn}			0.134^{***}	0.140^{***}
			(0.014)	(0.015)
$size_i$				0.009^{*}
				(0.005)
$D_i^{Tech} * \sigma_i^{eps}$		0.001	0.022^{**}	0.016^{*}
		(0.009)	(0.009)	(0.009)
$D_i^{Tech} * \sigma_i^{eps}$			-0.080^{***}	-0.094^{***}
			(0.021)	(0.022)
$D_i^{Tech} * size_i$				-0.018^{**}
				(0.007)
Constant	0.044^{***}	0.100^{***}	0.099^{***}	0.048
	(0.009)	(0.024)	(0.022)	(0.036)
Observations	580	580	580	580

Forecast covariance (γ_i) is my dependent variable and firms in the Technology sector (D_i^{Tech}) is my treatment variable. For each firm i, γ_i is the log of the covariance between projected and realized returns on stock i; D_i^{Tech} is a dummy that takes the value of one whenever the stock i belongs the CRSP Technology Index; σ_i^{eps} is the log of the standard deviation of changes in Earnings Per Share; σ_i^{retn} is the log of the standard deviation of stock returns. Variables related to the informational environment of firms are $size_i$, which corresponds to the log of the time-mean of market capitalization. From the 2,695 stocks in the sample, 290 correspond to the Technology index. In the matched sample, both the control and treatment groups are composed of 290 units. The number of unmatched observations is 2115. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

4.1. Discussion

Given the propensity score, the treatment is exogenous (Rosenbaum & Rubin, 1983). A balancing score is a function $p(\cdot)$ of the observed covariates $\boldsymbol{x}_i$, such that the conditional distribution of $\boldsymbol{x}_i$ given $p(\boldsymbol{x}_i)$ is the same for treated ($D_i^{Fin} = 1$) and control ($D_i^{Fin} = 0$) units. Formally, $\boldsymbol{x}_i \perp\!\!\!\perp D_i^{Fin} | p(\boldsymbol{x}_i)$. This requires that the treatment assignment is strongly ignorable. Treatment assignment is strongly ignorable if it is conditionally independent of the potential outcomes given $\boldsymbol{x}_i$ (unconfoundedness), and if $0 < p(D_i^{Fin} = 1 | \boldsymbol{x}_i) < 1$ (overlap). If this condition is met, given the vector of covariates $\boldsymbol{x}_i$, we have the condition:

$$(\gamma_{i,1}, \gamma_{i,0}) \quad \perp\!\!\!\perp \quad D_i^{Fin} | p(\boldsymbol{x}_i) \tag{4}$$

This outcome depends heavily on correct specification, and there is no guarantee of achieving perfect balance or eliminating bias. The condition of ignorability is challenging to test, and there is a possibility that it could fail completely.

In closing, much more research is needed on the topic of forecast quality. To capture non-linear relationships between the variables, testing various functional forms and allowing for interactions would deepen the understanding of how information uncertainty affects forecast quality. Bootstrapping techniques could address overfitting issues and add reliability to the results. An additional gain in reliability can be achieved by incorporating forecast quality metrics, such as mean absolute error or confidence intervals. Recently, De Silva & Thesmar (2024) show that the difference in mean squared error between subjective forecasters and the econometrician can be decomposed into three factors: noise, bias, and the informational advantage held by the forecasters.

5. Conclusions

Risk and information uncertainty are key components in portfolio management and shape investment decisions. Heightened information uncertainty makes it challenging for investors to make well-informed decisions, leading to lower future stock returns and larger investment mistakes. Stock valuations of sell-side analysts have the potential to reveal information on a firm's information uncertainty, when commercial incentives and their tendency toward optimistic forecasts are appropriately considered.

Analysts' forecasts on firms with higher information uncertainty exhibit a weaker covariance with realized values of stock prices. These findings also highlight analysts' ability to provide forecasts with informational content on more transparent firms, despite potential biases, reinforcing the idea that forecast quality offers a unique perspective on informational content. My findings are in line with the literature that exposes that information gathering in the stock market is costly, and provide an argument to understand why analysts, who provide biased forecasts, are compensated as information intermediaries. These results refine the understanding of the nuanced relationship between analyst reports, information uncertainty, and forecast quality, suggesting that forecast bias alone cannot fully capture the informational content of forecasts in the context of heightened information uncertainty. In addition, these suggest that further research is required in order to identify whether analyst coverage is a good proxy for the informational environment of firms, and provides an explanation for the mixed empirical results on the relation between analyst coverage and firm opacity.

Acknowledgements

I thank the anonymous referee for the careful review of my manuscript, as well as the helpful comments and valuable insights provided.

[Received: February 2024 — Accepted: October 2024]

References

Abarbanell, J. & Lehavy, R. (2003), 'Biased forecasts or biased earnings? the role of reported earnings in explaining apparent bias and over/underreaction in analysts' earnings forecasts', *Journal of Accounting and Economics* **36**(1-3), 105–146.

Aslan, H. & Kumar, P. (2017), 'Stapled financing, value certification, and lending efficiency', *Journal of Financial and Quantitative Analysis* **52**(2), 677–703.

Asquith, P., Mikhail, M. B. & Au, A. S. (2005), 'Information content of equity analyst reports', *Journal of financial economics* **75**(2), 245–282.

Banerji, S. & Basu, P. (2017), 'Universal banking, asymmetric information and the stock market', *Economic Modelling* **60**, 180–193.

Bartov, E., Faurel, L. & Mohanram, P. S. (2018), 'Can twitter help predict firm-level earnings and stock returns?', *The Accounting Review* **93**(3), 25–57.

Beaver, W. H. (2002), 'Perspectives on recent capital market research', *The Accounting Review* **77**(2), 453–474.

Benamar, H., Foucault, T. & Vega, C. (2021), 'Demand for information, uncertainty, and the response of us treasury securities to news', *The Review of Financial Studies* **34**(7), 3403–3455.

Beyer, A. & Guttman, I. (2011), 'The effect of trading volume on analysts' forecast bias', *The Accounting Review* **86**(2), 451–481.

Bilinski, P., Lyssimachou, D. & Walker, M. (2013), 'Target price accuracy: International evidence', *The Accounting Review* **88**(3), 825–851.

Blau, B. M., Brough, T. J. & Griffith, T. G. (2017), 'Bank opacity and the efficiency of stock prices', *Journal of Banking & Finance* **76**, 32–47.

Calomiris, C. W. & Kahn, C. M. (1991), 'The role of demandable debt in structuring optimal banking arrangements', *The American Economic Review* pp. 497–513.

Chan, K. & Hameed, A. (2006), 'Stock price synchronicity and analyst coverage in emerging markets', *Journal of Financial Economics* **80**(1), 115–147.

Cheng, Y., Liu, M. H. & Qian, J. (2006), 'Buy-side analysts, sell-side analysts, and investment decisions of money managers', *Journal of Financial and Quantitative Analysis* **41**(1), 51–83.

Corea, F. (2016), 'Can twitter proxy the investors' sentiment? the case for the technology sector', *Big Data Research* **4**, 70–74.

Cowen, A., Groysberg, B. & Healy, P. (2006), 'Which types of analyst firms are more optimistic?', *Journal of Accounting and Economics* **41**, 119–146.

Crawford, S. S., Roulstone, D. T. & So, E. C. (2012), 'Analyst initiations of coverage and stock return synchronicity', *The Accounting Review* **87**(5), 1527–1553.

Dahiya, S., Iannotta, G. & Navone, M. (2017), 'Firm opacity lies in the eye of the beholder', *Financial Management* **46**(3), 553–592.

De Silva, T. & Thesmar, D. (2024), 'Noise in expectations: Evidence from analyst forecasts', *The Review of Financial Studies* **37**(5), 1494–1537.

DeMiguel, V. & Nogales, F. J. (2009), 'Portfolio selection with robust estimation', *Operations Research* **57**(3), 560–577.

Derrien, F. & Kecskes, A. (2013), 'The real effects of financial shocks: Evidence from exogenous changes in analyst coverage', *The Journal of Finance* **68**(4), 1407–1440.

Fischer, P. E. & Stocken, P. C. (2010), 'Analyst information acquisition and communication', *The Accounting Review* **85**(6), 1985–2009.

Flannery, M. J., Kwan, S. H. & Nimalendran, M. (2004), 'Market evidence on the opaqueness of banking firms' assets', *Journal of Financial Economics* **71**(3), 419–460.

Forbes, W. P., Murphy, A., O'Keeffe, C. & Su, C. (2020), 'Are financial analysts eager postmen of bubble psychology? evidence in the united kingdom', *International Journal of Finance & Economics* **25**(1), 120–137.

Greenwood, B. N. & Gopal, A. (2015), 'Research note - tigerblood: Newspapers, blogs, and the founding of information technology firms', *Information Systems Research* **26**(4), 812–828.

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

Hansen, L. P. & Sargent, T. J. (2010), 'Fragile beliefs and the price of uncertainty', *Quantitative Economics* **1**(1), 129–162.

Hao, X., Wang, Y., Wu, C. & Wu, L. (2024), 'Oil information uncertainty and aggregate market returns: A natural experiment based on satellite data', *Journal of Financial Markets* p. 100913.

Hilary, G. & Hsu, C. (2013), 'Analyst forecast consistency', *The Journal of Finance* **68**(1), 271–297.

Huang, A. H., Zang, A. Y. & Zheng, R. (2014), 'Evidence on the information content of text in analyst reports', *The Accounting Review* **89**(6), 2151–2180.

Huang, R. & Ratnovski, L. (2011), 'The dark side of bank wholesale funding', *Journal of Financial Intermediation* **20**(2), 248–263.

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

Jia, N., Rai, A. & Xu, S. X. (2020), 'Reducing capital market anomaly: The role of information technology using an information uncertainty lens', *Management Science* **66**(2), 979–1001.

Jiang, G., Lee, C. M. & Zhang, Y. (2005), 'Information uncertainty and expected returns', *Review of Accounting Studies* **10**, 185–221.

Jones, J. S., Lee, W. Y. & Yeager, T. J. (2013), 'Valuation and systemic risk consequences of bank opacity', *Journal of Banking & Finance* **37**(3), 693–706.

Joos, P., Piotroski, J. D. & Srinivasan, S. (2016), 'Can analysts assess fundamental risk and valuation uncertainty? an empirical analysis of scenario-based value estimates', *Journal of Financial Economics* **121**(3), 645–663.

Kang, W., Li, N. & Zhang, H. (2019), 'Information uncertainty and the pricing of liquidity', *Journal of Empirical Finance* **54**, 77–96.

Kumar, A. (2009), 'Hard-to-value stocks, behavioral biases, and informed trading', *Journal of Financial and Quantitative Analysis* **44**(6), 1375–1401.

Lim, T. (2001), 'Rationality and analysts' forecast bias', *The Journal of Finance* **56**(1), 369–385.

Lys, T. & Soo, L. G. (1995), 'Analysts' forecast precision as a response to competition', *Journal of Accounting, Auditing & Finance* **10**(4), 751–765.

Maccheroni, F., Marinacci, M. & Ruffino, D. (2013), 'Alpha as ambiguity: Robust mean-variance portfolio analysis', *Econometrica* **81**(3), 1075–1113.

Malmendier, U. & Shanthikumar, D. (2014), 'Do security analysts speak in two tongues?', *The Review of Financial Studies* **27**(5), 1287–1322.

Matutes, C. & Vives, X. (1996), 'Competition for deposits, fragility, and insurance', *Journal of Financial Intermediation* **5**(2), 184–216.

Mehran, H. & Peristiani, S. (2009), 'Financial visibility and the decision to go private', *The Review of Financial Studies* **23**(2), 519–547.

Mikhail, M. B., Walther, B. R. & Willis, R. H. (1999), 'Does forecast accuracy matter to security analysts?', *The Accounting Review* **74**(2), 185–200.

Morgan, D. P. (2002), 'Rating banks: Risk and uncertainty in an opaque industry', *American Economic Review* **92**(4), 874–888.

Piotroski, J. D. & Roulstone, D. T. (2004), 'The influence of analysts, institutional investors, and insiders on the incorporation of market, industry, and firm-specific information into stock prices', *The Accounting Review* **79**(4), 1119–1151.

Quandt, R. E. (1958), 'The estimation of the parameters of a linear regression system obeying two separate regimes', *Journal of the american statistical association* **53**(284), 873–880.

Quandt, R. E. (1972), 'A new approach to estimating switching regressions', *Journal of the American statistical association* **67**(338), 306–310.

Rosenbaum, P. R. & Rubin, D. B. (1983), 'The central role of the propensity score in observational studies for causal effects', *Biometrika* **70**(1), 41–55.

Sambhara, C., Rai, A. & Xu, S. X. (2022), 'Configuring the enterprise systems portfolio: The role of information risk', *Information Systems Research* **33**(2), 446–463.

Su, C., Zhang, H., Bangassa, K. & Joseph, N. L. (2019), 'On the investment value of sell-side analyst recommendation revisions in the uk', *Review of Quantitative Finance and Accounting* **53**, 257–293.

Tao, R., Brooks, C. & Bell, A. R. (2020), 'When is a max not the max? how news resolves information uncertainty', *Journal of Empirical Finance* **57**, 33–51.

Wieland, M. M. (2011), 'Identifying consensus analysts' earnings forecasts that correctly and incorrectly predict an earnings increase', *Journal of Business Finance & Accounting* **38**(5-6), 574–600.

Zhang, X. F. (2006), 'Information uncertainty and stock returns', *Journal of Finance* **61**, 105–136.