Impact of FINRA 2241

Rajeev Bhattacharya (✉️ RRB@Washington-Finance.com)
Washington Finance and Economics

Research Article

Keywords: Market Efficiency, Event Studies, Analyst Forecasts, FINRA 2241

Posted Date: October 28th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2196685/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Impact of FINRA 2241

Rajeev R. Bhattacharya*

October 22, 2022

Abstract

I calculate six separate measures of market efficiency based on event studies using intraday data on equities and fixed income securities. I calculate diligence, objectivity, quality of analysis, and accuracy, of analyst forecasts. The FINRA 2241 Rules (2015) were the first major attempt by a U.S. self-regulatory organization to specifically enforce diligence and objectivity by analysts. I model equations that have a) each metric as a function of exogenous factors, namely, indicator variable for post-FINRA 2241, Nasdaq listing, dispersion in investor valuations, short sales costs & constraints, and transaction costs & constraints, and endogenous market activities, namely, Kyle-Obizhaeva liquidity measure, normalized short interest, analyst coverage, institutional ownership of equity, log of market cap, and log of shares outstanding, and b) each endogenous market activity as a function of the exogenous factors and all other endogenous market activities. I use the panel nature of the data to identify appropriate instruments for the endogenous variables and for the variables that are measured with error, and I use Three Stage Least Squares and Errors in Variables to estimate this seven-equation structural model. I find that the impact of FINRA 2241 on each of these ten objective and systematic market quality metrics is economically insignificant.

Keywords: Market Efficiency; Event Studies; Analyst Forecasts; FINRA 2241.

JEL Codes: G12; G14; G24; C23; C26; K22.

Statements and Declarations.

No funds, grants, or other support was received. I have no financial, proprietary, or non-financial interests in any material discussed in this article. I have nothing to disclose.

Acknowledgments.

I acknowledge, with gratitude, the invaluable thoughts provided by Reena Aggarwal, Prerana Agrawal, Matt Armstrong, Ray Ball, Sudipta Basu, Jeremy Bertomeu, Jacqueline Birt, Cynthia Cai, Jivas Chakravarthy, Ryan Cooper, Martijn Cremers, Ted Davis, Ahmed El Nahas, Philip Evans, Mark Flood, Hila Fogel-Yaari, Simon Grant, Margaret Hansen, Gautam Hazarika, David Hirshleifer, Denada Ibrushi, Barclay James, Siamak Jawadi, Gergana Jostova, Ankit Kapur, Wanzhi Ke, Hwee Ping Koh, Leo Langa, Hong-Hee Lee, John Lee, Julie Lee, Andi Li, Zecong Ma, Marcie Mao, Stijn Masschelein, Michael McDonald, Andre Mollick, Vikram Nanda, Nam Hoang Nguyen, Lee Ohanian, Prasad Padmanabhan, Jeffrey Pontiff, Tava Olsen, Elizabeth Ooi, Terry Pan, Jeff Pontiff, Nagpurnanand Prabhala, Sanjiv Sabherwal, Hanlin Shang, Hersh Shefrin, Lei Shi, Erik Sirri, Pablo Slutzky, Bin Srinidhi, Raghu Sundaram, Kien Tran, Aadhaar Verma, David Weinbaum, Simon Wheatley, Yi Xiang, Yexiao Xu, and Hamid Yahyaei (in alphabetical order), and especially, Henk Berkman, Philip Brown, Mahendra Gupta, Robert Korajczyk, Pete Kyle, Glenn MacDonald, Tom Smith, Lei Shi, and seminar participants at Australian National University, University of Auckland, Macquarie University, University of Maryland at College Park, St. Mary’s University, University of Texas at Arlington, University of Texas at Dallas, University of Texas at Rio Grande Valley, University of Warsaw, University of Western Australia, and the U.S. Securities and Exchange Commission. The usual caveats apply.
1 Introduction

Bhattacharya (2022) uses two different metrics — based on a) abnormal responses to “key developments” (as indicated by Standard & Poor’s) and b) earnings announcements and revisions, and analyst forecasts and revisions — with three separate 60-minute, 90-minute, and 120-minute announcement windows for each, based on event studies, controlling for market equity returns, Nasdaq listing equity returns, industry (3-digit NAICS Code) equity returns, market cap decile equity returns, intraday volatility decile equity returns, dividend decile equity returns, Fama-French factor decile equity returns, fixed income yield, and foreign exchange, with intraday equity and fixed income securities data on all publicly traded U.S. companies over 2014-September 2018 (tens of trillions of observations,\(^1\) about 30 TB of data\(^2\)), as separate objective and systematic ordinal measures of the efficiency of the market for a stock.\(^3\) It uses the market microstructure models in Kyle and Obizhaeva (2016) and Bhattacharya (2019-b) to develop a seven-equation structural model.

Bhattacharya and Gupta (2022) provides a general framework of behavior under asymmetric information and applies this framework to an analyst’s forecasts about a studied firm as a function of publicly available information and the use and processing of non-public information, when the studied firm provides earnings guidance to the market as well.\(^4\) Diligence by an analyst-analyst-firm about a studied firm would require the analyst-analyst-firm acquiring and analyzing non-public information specific to the studied firm, even potentially information that is not available to the studied firm itself. Comparison of the Public Model versus the General Model provides the index of non-public information by the analyst-analyst-firm about the studied firm, in this framework, the index of private information equals one minus the \(p\)-value of the Hausman Specification Test of Ordinary Least Squares (OLS) versus Two Stage Least Squares (2SLS) of earnings forecast on

---

\(^1\)For perspective, the total number of stars in the Milky Way is estimated to be 100 billion, less than a hundredth of the number of observations analyzed in this paper.

\(^2\)Analyzing Big Data is not a scalable version of the programming that is done for smaller datasets.

\(^3\)Bhattacharya (2022-b) uses a similar framework for efficiency of the market for a fixed income security.

\(^4\)I use multiplicative models in this paper to make my models comparable across firms and industries, and across time. In order to implement the multiplicative models that make comparisons possible, I restrict ourselves to positive earnings, management guidance, and analyst forecasts. Please see Bhattacharya and Gupta (2022) for the implications of this restriction.
management guidance and public information.\textsuperscript{5} Objectivity by an analyst-analyst-firm about a studied firm would require the analyst-analyst-firm’s forecast accurately reporting its best estimate about the studied firm’s earnings. Comparison of the Objectivity Model versus the general Non-Objectivity Model provides the index of objectivity by the analyst-analyst-firm about the studied firm, which equals the $p$-value of the Wald Test of zero coefficients versus non-zero coefficients in the 2SLS regression of the earnings forecast residual on publicly available information. The exponent of the negative of the standard deviation of residuals of the analyst forecast regression equation provides the index of analytical quality by the analyst-analyst-firm about the studied firm.\textsuperscript{6} Under the assumption that the announced earnings of studied firm $i$ for quarter $t$, announced to the public at the beginning of quarter $(t + 1)$, are unbiased for actual earnings of the studied firm for quarter $t$, the index of \textit{ex post} normalized accuracy for period $t$ by the analyst-analyst-firm about the studied firm is given by the exponent of the negative of the absolute difference between log of analyst forecast for period $t$ and log of announced earnings for period $t$, when both are positive. They find that \textit{ex post} normalized accuracy by an analyst-analyst-firm about a studied firm is a statistically and economically significant increasing function of the product of the indices of diligence, objectivity, and quality by the analyst-analyst-firm about the studied firm.

Please see Table 1 for summary statistics for these ten separate ordinal market quality metrics.

The FINRA\textsuperscript{7} 2241 Rules (effective September 25, 2015, and December 24, 2015) were the first major attempt by a U.S. self-regulatory organization (SRO) to specifically enforce diligence and objectivity by analysts. “A member’s written policies and procedures must be reasonably designed to promote objective and reliable research that reflects the truly held opinions of research analysts and to prevent the use of research reports or research analysts to manipulate or condition the market or favor the interests of the member or a current or prospective customer or class of customers. ... A member must establish, maintain and enforce written policies and procedures reasonably designed to ensure that: A) purported facts in its research reports are based on reliable information; and (B) any recommendation, rating or price target has a reasonable basis and is accompanied by a clear

\textsuperscript{5}See, for instance, Hausman (1978) and Greene (2018).
\textsuperscript{6}See, for instance, Greene (2018).
\textsuperscript{7}Financial Industry Regulatory Authority (https://www.finra.org/about).
explanation of any valuation method used and a fair presentation of the risks that may impede achievement of the recommendation, rating or price target."8

In this paper, I apply the methodology of Bhattacharya (2022) to model equations that have a) each market quality metric as a function of exogenous factors, namely, indicator variable for post-FINRA 2241, Nasdaq listing, dispersion in investor valuations (proxied by the quarterly standard deviation of analyst forecasts of earnings per share for the firm with “forecast period end date” in the current calendar quarter), short sales costs & constraints (proxied by the cost of a synthetic short sale, which is short one call option and long one put option, at the money, with the same expiration), and transaction costs & constraints (proxied by bid-ask spread), and endogenous market activities, namely, Kyle-Obizhaeva liquidity measure,9 normalized short interest, analyst coverage, institutional ownership of equity, log of market cap, and log of shares outstanding, and b) each endogenous market activity as a function of the exogenous factors and all other endogenous market activities. I use the panel nature of the data to identify appropriate instruments for the endogenous variables and for the variables that are measured with error, and I use Three Stage Least Squares (3SLS) and Errors in Variables (EiV) to estimate this seven-equation structural model and test the hypotheses provided by the theory. I develop the methodology to calculate the total impact of a factor which is the sum of the direct impact and the indirect impacts, and by using the Gaussian cumulative distribution of the $Z$-score of each variable (except for an indicator variable), I make the impacts comparable. Using an indicator variable which is zero for quarters prior to Q3-2015 and one for quarters after Q4-2015, and deleting observations for Q3-2015 and Q4-2015, I find that FINRA 2241’s impact on each of the above ten objective and systematic market quality metrics was

9See Kyle and Obizhaeva (2016).
The paper is organized as follows. Section 2 describes the data. Section 3 details the methodology of this paper. Section 4 provides the main results. Section 5 concludes.

2 Description of Data

I use data on all publicly traded U.S. stocks for 2014-September 2018, please see Law (2021) regarding “anonymization” and “reshuffling” of analyst and analyst firm codes in I/B/E/S data, as a result of which I/B/E/S data beyond October 2018 are unreliable.\(^\text{11}\)

I use intraday equity trading data from TAQ and intraday fixed income security trading data from TRACE. I divide each trading day into 15 bins as follows: bin 0 for prior to 9:30 AM, bin 14 for after 4 PM, and bins 1-13 for each half-hour of the trading hours. For each stock, for each bin 0-14 with positive volume, I calculate the volume-weighted average price (VWAP) of trading prices, and then calculate the relevant bin-to-bin continuously compounded return for each stock. For each fixed income (FI) security CUSIP for each bin, I calculate the volume-weighted average yield (VWAY) of all FI trades,\(^\text{12}\) then for each TICKER, I calculate the simple average of VWAY over all CUSIPs corresponding to that TICKER.

I restrict attention to stocks corresponding to publicly traded firms that did not change PERMNO or TICKER or 3-digit NAICS code during 2014-September 2018. I further restrict attention to stock-days that did not have splits, reverse splits, or dividends, and to stock-days\(^\text{13}\) that have positive closing price and positive shares outstanding\(^\text{14}\) from CRSP: date, closing price, return, shares outstanding, trading volume, closing bid and ask, exchange membership, NAICS code, number of analysts, short interest, and institutional ownership — not as a zero.\(^\text{13}\)

\(^{10}\) I have repeatedly engaged with the team of economists at FINRA to address these issues, but have not received a meaningful response.

\(^{11}\) I thank Henk Berkman for the invaluable contribution of pointing this out.

\(^{12}\) With \(|y|\leq 100\%\).

\(^{13}\) In order to maintain the integrity of the linkages between the different sources of data I rely upon, I treat a missing observation as a missing observation — in particular, with number of analysts, short interest, and institutional ownership — not as a zero.

\(^{14}\) For a small number of stock-days, I found multiple entries when sorted by ticker, CUSIP, and date — for these stock-days, I calculate weighted averages using trading volume as weights.
and had data on the following variables: analyst forecasts and revisions (from I/B/E/S), earnings announcements and revisions (from I/B/E/S),\textsuperscript{15} institutional holdings (from Thomson Reuters), short interest (from Compustat),\textsuperscript{16} Key Developments (from Capital IQ), Compustat-CRSP Merged Database,\textsuperscript{17} options data from OptionMarket Quality Metrics, and Nominal Broad U.S. Dollar Index from Federal Reserve of St. Louis (FRED). Since FINRA 2241 was implemented in September and December 2015, I delete observations for Q3-2015 and Q4-2015.

These data constitute about 30 TB, tens of trillions of observations, are available at https://drive.google.com/drive/folders/107COBM5_jZ_Fli1dYPH0bngvquMWvm5H?usp=sharing.

3 Methodology

Based on the market microstructure models in Kyle and Obizhaeva (2016) and Bhattacharya (2019-b), Bhattacharya (2022) develops a seven-equation structural system and estimates it at the stock-quarter level, where \( i \) represents the stock and \( t \) represents the quarter.

3.1 Notations

- \( Metric_{i,t} \) = each separate ordinal market quality metric for stock \( i \) in quarter \( t \), described in Section 1
- \( FINRA2241_t \) = indicator variable which is zero for quarters prior to Q3-2015 and one for quarters after Q4-2015, and deleting observations for Q3-2015 and Q4-2015
- \( Nasdaq_{i,t} \) = indicator variable for Nasdaq listing of firm \( i \)
- \( Disp_{i,t} \) = dispersion in investor valuations for stock \( i \) in quarter \( t \)
- \( SSCC_{i,t} \) = short sales costs and constraints of stock \( i \) in quarter \( t \)

\textsuperscript{15}With \(|EPS| \leq $100. 
\textsuperscript{16}I do not have access to actual short sales costs & constraints data — dealing with missing data is one of the biggest challenges in empirical work, and appropriately proxying for the unavailable data is one of the innovations of this paper.
\textsuperscript{17}Compustat (including Capital IQ) data sources rely upon the S&P identification of GVKEY. In this paper, I rely upon the mapping of GVKEY to TICKER used by CRSP to match its data to Compustat.
• $TrCC_{i,t} = \text{transaction costs & constraints of stock } i \text{ in quarter } t$

• $KOLiq_{i,t} = \text{Kyle-Obizhaeva liquidity of stock } i \text{ in quarter } t = \text{cube root (quarterly mean of bin dollar volume)/(square of quarterly standard deviation of bin returns)}^{18}$

• $NormSI_{i,t} = \text{quarterly mean of (monthly short interest)/(daily shares outstanding) of stock } i \text{ in quarter } t$

• $AnCov_{i,t} = \text{number of EPS forecasts and revisions by analysts for firm } i \text{ with “forecast period end date” in quarter } t$

• $NormInst_{i,t} = \text{quarterly mean of institutional ownership percentage in stock } i \text{ in quarter } t$

• $logMCap_{i,t} = \text{quarterly mean of natural logarithm of daily market cap for stock } i \text{ in quarter } t$

• $logShrsOut_{i,t} = \text{quarterly mean of natural logarithm of daily shares outstanding for stock } i \text{ in quarter } t$

3.2 Transaction Costs & Constraints

I use normalized bid-ask spread, the quarterly mean of daily relative bid-ask spread where daily relative bid-ask spread = (daily closing best ask – daily closing best bid)/(mid-point of daily closing best bid and best ask), as a proxy for transaction costs — see, for instance, Hasbrouck (2009).

3.3 Short Sales Costs & Constraints

I do not have access to actual data on short sales costs & constraints.\textsuperscript{19,20} Dealing with missing or unavailable data is one of the biggest challenges in empirical work, and appropriately proxying for the unavailable data is one of the innovations of this paper. In particular, I use the cost of a

\textsuperscript{18}See Kyle and Obizhaeva (2016).
\textsuperscript{19}I have tried many times, with very influential deans of business schools, as well as C.E.O.s of investment banks, but since the Great Recession, such data have become essentially unavailable to a researcher.
\textsuperscript{20}See Ofek, Richardson and Whitelaw (2004) and Cremers and Weinbaum (2010), for instance, on papers that use proprietary data on rebate rates and spreads.
synthetic short sale (short one call option and long one put option, at the money, with the same expiration) as a measure of short sales costs & constraints for the underlying stock — please see, in reverse chronological order, Evans, Geczy, Musto and Reed (2009), Lamont and Thaler (2003), and Geczy, Musto and Reed (2002) — as follows:

• Consider all options with strike prices within $+/- 2.5\%$ of underlying price$^{21}$

• Calculate best ask for put minus best bid for call

• Restrict to positive

• Calculate weighted (by volume) average

• Normalize by underlying price

3.4 Dispersions of Investor Valuations

The dispersion of investor valuations, which is also a function of the level of uncertainty within the market about the particular firm $i$ in quarter $t$, is proxied by calculating the quarterly standard deviation of analyst forecasts and analyst forecast revisions of earnings per share (EPS) for firm $i$, restricted to forecasts and revisions between $-$100 per share and +$100 per share, for the firm $i$ with each “forecast period end date” in the current calendar quarter $t$.

$^{21}$I do sensitivity analyses for $+/- 0.5\%$ and $+/- 5\%$, and the results stay qualitatively the same.
3.5 Market Quality Metrics

Each market quality metric is modeled as a linear function of the exogenous factors and endogenous market activities as follows.

\[
Metric_{i,t} = \gamma + \gamma_{FINRA2241}FINRA2241_t + \gamma_{Nasdaq}Nasdaq_{i,t} + \gamma_{Disp}Disp_{i,t} + \gamma_{SSCC}SSCC_{i,t} + \gamma_{TrCC}TrCC_{i,t} + \gamma_{KOLiq}KOLiq_{i,t} + \gamma_{NormSI}NormSI_{i,t} + \gamma_{AnCov}AnCov_{i,t} + \gamma_{NormInst}NormInst_{i,t} + \gamma_{logMCap}logMCap_{i,t} + \gamma_{logShrsOut}logShrsOut_{i,t} + \delta_{i,t}
\]

3.6 Endogenous Market Activities

Each endogenous market activity is modeled as a linear function of exogenous factors and other endogenous market activities, based on the market microstructure models in Kyle and Obizhaeva (2016) and Bhattacharya (2019-b), leading to the following six equations.

\[
KOLiq_{i,t} = \alpha + \alpha_{FINRA2241}FINRA2241_t + \alpha_{Nasdaq}Nasdaq_{i,t} + \alpha_{Disp}Disp_{i,t} + \alpha_{SSCC}SSCC_{i,t} + \alpha_{TrCC}TrCC_{i,t} + \alpha_{NormSI}NormSI_{i,t} + \alpha_{AnCov}AnCov_{i,t} + \alpha_{NormInst}NormInst_{i,t} + \alpha_{logMCap}logMCap_{i,t} + \alpha_{logShrsOut}logShrsOut_{i,t} + \varepsilon_{i,t}
\]
\[ \text{Norm}_{SI,t} = \theta + \theta_{FINRA2241} \text{FINRA2241}_{t} + \theta_{Nasdaq} \text{Nasdaq}_{i,t} + \theta_{Disp} \text{Disp}_{i,t} + \theta_{SSCC} \text{SSCC}_{i,t} + \theta_{TrCC} \text{TrCC}_{i,t} + \theta_{KOLiq} \text{KOLiq}_{i,t} + \theta_{NormInst} \text{NormSI}_{i,t} + \theta_{AnCov} \text{AnCov}_{i,t} + \theta_{logMCap} \text{logMCap}_{i,t} + \theta_{logShrsOut} \text{logShrsOut}_{i,t} + \eta_{i,t} \]

\[ \text{AnCov}_{i,t} = \zeta + \zeta_{FINRA2241} \text{FINRA2241}_{i,t} + \zeta_{Nasdaq} \text{Nasdaq}_{i,t} + \zeta_{Disp} \text{Disp}_{i,t} + \zeta_{SSCC} \text{SSCC}_{i,t} + \zeta_{TrCC} \text{TrCC}_{i,t} + \zeta_{KOLiq} \text{KOLiq}_{i,t} + \zeta_{NormInst} \text{NormSI}_{i,t} + \zeta_{NormInst} \text{NormInst}_{i,t} + \zeta_{logMCap} \text{logMCap}_{i,t} + \zeta_{logShrsOut} \text{logShrsOut}_{i,t} + \theta_{i,t} \]

\[ \text{NormInst}_{i,t} = \psi + \psi_{FINRA2241} \text{FINRA2241}_{t} + \psi_{Nasdaq} \text{Nasdaq}_{i,t} + \psi_{Disp} \text{Disp}_{i,t} + \psi_{SSCC} \text{SSCC}_{i,t} + \psi_{TrCC} \text{TrCC}_{i,t} + \psi_{KOLiq} \text{KOLiq}_{i,t} + \psi_{NormInst} \text{NormSI}_{i,t} + \psi_{AnCov} \text{AnCov}_{i,t} + \psi_{logMCap} \text{logMCap}_{i,t} + \psi_{logShrsOut} \text{logShrsOut}_{i,t} + \nu_{i,t} \]

\[ \text{logMCap}_{i,t} = \xi + \xi_{FINRA2241} \text{FINRA2241}_{t} + \xi_{Nasdaq} \text{Nasdaq}_{i,t} + \xi_{TrCC} \text{TrCC}_{i,t} + \xi_{SSCC} \text{SSCC}_{i,t} + \xi_{Disp} \text{Disp}_{i,t} + \xi_{KOLiq} \text{KOLiq}_{i,t} + \xi_{NormInst} \text{NormSI}_{i,t} + \xi_{AnCov} \text{AnCov}_{i,t} + \xi_{NormInst} \text{NormInst}_{i,t} + \xi_{logShrsOut} \text{logShrsOut}_{i,t} + \nu_{i,t} \]
\[ \logShrsOut_{i,t} = \omega + \omega_{FINRA2241} \text{FINRA2241}_t + \omega_{Nasdaq}\text{Nasdaq}_{i,t} \\
+ \omega_{Disp}\text{Disp}_{i,t} + \omega_{SSCC}\text{SSCC}_{i,t} + \omega_{TrCC}\text{TrCC}_{i,t} \\
+ \omega_{KOLiq}\text{KOLiq}_{i,t} + \omega_{NormInst}\text{NormInst}_{i,t} + \omega_{AnCov}\text{AnCov}_{i,t} \\
+ \omega_{logMCap}\text{logMCap}_{i,t} + \varrho_{i,t} \]

3.7 Summary Statistics

Table 1
Correlation Coefficients of All Ordinal Market Quality Metrics
At Year-Quarter Level (2014-September 2018)

Table 2.1
Summary Statistics of All Potential Factors and Market Activities
At Year-Quarter Level (2014-September 2018)

Table 2.2
Correlation Coefficients of All Potential Factors and Market Activities
At Year-Quarter Level (2014-September 2018)

I use Three Stage Least Squares (3SLS) and Errors in Variables (EiV) to account for the endogeneities and simultaneities in my model.\(^{22}\)

3.8 Instruments

In a system of equations \( Y = X\beta + \varepsilon \), for a vector \( W \) to be an instrument, it is required that **Strong**
First Stage: \( \text{Cov}(W, X) \neq 0 \), and **Exclusion Restriction**: \( \text{Cov}(W, \varepsilon) = 0.\)\(^{23}\) As Wooldridge

\(^{22}\) Please see Liu and Saraiva (2019), Kahouli (2018), and Lee, Liang, Lin and Yang (2016), in reverse chronological order, for robustness of this methodology.

In particular, “under conditional homoskedasticity, GMM reduces to 3SLS if the set of instrumental variables is common to all equations.” Kahouli (2018).

\(^{23}\) In other words, the unexplained part (“residual”) \( \varepsilon \) of the regression needs to be uncorrelated with the instrument — this *does not* require that the regressand \( Y \) itself be uncorrelated with the instrument.
(2010) points out, most of the problems with using instrumental variables arise in small samples, which, obviously, is not the case in this paper. Wooldridge (2010) points out that “asymptotically, we can do no worse, and can often do better, using a larger set of valid instruments.” Given these two critical qualifications, I consider in this paper the following instrumentation categories together to ensure that my conclusions are not sensitive to particular information assumptions. In panel data econometrics, there are two popular categories of autochthonous instruments.

1. **Cross Sectional Instruments.** Instrument a variable $X_{i,t}$ by combination(s) of $X_{j,t}, j \neq i$. Cross sectional variables (such as corresponding variables for other geographies) have been used as instruments by, for instance, Hausman, Leonard and Zona (1994) and the validity of such instrumentation depends on the assumption that the corresponding cross sectional variables are correlated with the instrumented variable because of cost (or other) commonalities in the time period of interest but are not influenced by the idiosyncrasies of the stock (or geography) of interest in the particular time period. In particular, I use, as cross sectional instrument the average across the cross section excluding the particular variable, i.e., for variable $X_{i,t}$, I use $\text{average}_{j\neq i}(X_{j,t})$ as instrument.

2. **Time Series Instruments.** Instrument a variable $X_{i,t}$ by its lag(s) $X_{i,t-1}, X_{i,t-2}, ...$ The use of lags as instruments depends on two implicit assumptions: a) lack of serial correlation, see, for example, Greene (2018), and b) weak rational expectations, see, for example, Hansen and Singleton (1982). In particular, I use as time series instrument the first lag of the relevant variable, i.e., for variable $X_{i,t}$, I use $X_{i,t-1}$ as instrument.

---

24Wintoki, Linck and Netter (2012) and Roberts and Whited (2011) show that the traditional statistical tests for the use of lagged variables in panel estimation are not particularly useful, so I do not use them in this paper.
3.9 Z-Score, Cumulative Distribution Function, and Economic Significance

For ease of interpretation, I replace each variable \( x \), except for indicator and time variables, by its normalized version

\[
\Phi(Z_{\text{Score}}(x)) = \Phi\left(\frac{x - \text{Mean}(x)}{\text{StDev}(x)}\right)
\]

where \( \Phi \) is the cumulative distribution function of a standard Gaussian random variable.

Therefore, all such regression coefficients are comparable, a coefficient \( \beta(y, x) \) on the regression of the regressand \( y \) on the regressor \( x \) means that a 1% increase in the cumulative probability of \( x \) causes a \( \beta(y, x) \)% increase in the cumulative probability of \( y \).

Similarly, a regression coefficient \( \beta(y, I) \) on the regression of the regressand \( y \) on the indicator variable \( I \) means that there is a \( \beta(y, I) \) higher cumulative probability of \( y \).

The impact of a regressor \( x \) on the regressand \( y \) is **economically significant positive** if the relevant coefficient\(^{26}\)

\[
\beta(y, x) > 0.01
\]

and is **economically significant negative** if the relevant coefficient

\[
\beta(y, x) < -0.01
\]

I indicate economically significant positive impacts by green highlighting and economically significant negative impacts by red highlighting.

\(^{25}\)It is worth pointing out that \( \text{Percentile}(x) \approx \text{RoundUp}(100\Phi(Z_{\text{Score}}(x))) \) and, therefore, \( |\beta(y, x)| \geq 1 \) implies that a move of \( x \) to one higher percentile causes \( y \) to move up (approximately) \( \text{Round}(\beta(y, x)) \) percentiles, and \( 0 \leq |\beta(y, x)| < 1 \) implies that a move of \( x \) to one higher percentile causes \( y \) to stay in (approximately) the same percentile. Similarly, \( \text{Millenile}(x) \approx \text{RoundUp}(1,000\Phi(Z_{\text{Score}}(x))) \), \( \text{Decile}(x) \approx \text{RoundUp}(10\Phi(Z_{\text{Score}}(x))) \), \( \text{Quartile}(x) \approx \text{RoundUp}(4\Phi(Z_{\text{Score}}(x))) \), etc.

\(^{26}\)Please see Aman, et. al., (2019) for a robust defense of “responsible science,” that science needs to have integrity and relevance.
4 Total Impact of Each Exogenous Factor on Each Market Quality Metric

Given a differentiable function \( F(x_1, x_2, ..., x_m, y_1, ..., y_n) : \mathbb{R}^{m+n} \rightarrow \mathbb{R} \), where \( y_j = y_j(x_1, x_2, ..., x_m), j = 1, ..., n \),

\[
\frac{dF(x_1, x_2, ..., x_m, y_1, ..., y_n)}{dz} = \frac{\partial F}{\partial x_1} \frac{dx_1}{dz} + \frac{\partial F}{\partial x_2} \frac{dx_2}{dz} + ... + \frac{\partial F}{\partial x_m} \frac{dx_m}{dz} \\
+ \frac{\partial F}{\partial y_1} \frac{dy_1}{dz} + \frac{\partial F}{\partial y_2} \frac{dy_2}{dz} + ... + \frac{\partial F}{\partial y_n} \frac{dy_n}{dz}
\]

and thus, \( \forall k = 1, ..., m \),

\[
\frac{dF}{dx_k} \text{ (total impact)} = \frac{\partial F}{\partial x_k} \text{ (direct impact)} \\
+ \frac{\partial F}{\partial y_1} \frac{dy_1}{dx_k} + \frac{\partial F}{\partial y_2} \frac{dy_2}{dx_k} + ... + \frac{\partial F}{\partial y_n} \frac{dy_n}{dx_k} \\
\text{(indirect impact)}
\]

And, therefore, the total impact of each exogenous factor on each market quality metric is given by the corresponding (cross-equation) total derivative as follows.
\[
\frac{d\text{Metric}}{d\text{FINRA2241}} = \frac{\partial\text{Metric}}{\partial\text{FINRA2241}} + \frac{\partial\text{Metric}}{\partial \text{KOLiq}} \frac{d\text{KOLiq}}{d\text{FINRA2241}} + \frac{\partial\text{Metric}}{\partial \text{NormSI}} \frac{d\text{NormSI}}{d\text{FINRA2241}} \\
+ \frac{\partial\text{Metric}}{\partial \text{AnCov}} \frac{d\text{AnCov}}{d\text{FINRA2241}} + \frac{\partial\text{Metric}}{\partial \text{NormInst}} \frac{d\text{NormInst}}{d\text{FINRA2241}} \\
+ \frac{\partial\text{Metric}}{\partial \text{logMCap}} \frac{d\text{logMCap}}{d\text{FINRA2241}} \frac{d\text{logShrsOut}}{d\text{FINRA2241}}
\]

\[
= \gamma_{\text{FINRA2241}} \\
+ \gamma_{\text{KOLiq}} \alpha_{\text{FINRA2241}} + \gamma_{\text{NormSI}} \theta_{\text{FINRA2241}} \\
+ \gamma_{\text{AnCov}} \zeta_{\text{FINRA2241}} + \gamma_{\text{NormInst}} \psi_{\text{FINRA2241}} \\
+ \gamma_{\text{logMCap}} \xi_{\text{FINRA2241}} + \gamma_{\text{logShrsOut}} \omega_{\text{FINRA2241}}
\]

\[
\frac{d\text{Metric}}{d\text{Nasdaq}} = \frac{\partial\text{Metric}}{\partial\text{Nasdaq}} + \frac{\partial\text{Metric}}{\partial \text{KOLiq}} \frac{d\text{KOLiq}}{d\text{Nasdaq}} + \frac{\partial\text{Metric}}{\partial \text{NormSI}} \frac{d\text{NormSI}}{d\text{Nasdaq}} \\
+ \frac{\partial\text{Metric}}{\partial \text{AnCov}} \frac{d\text{AnCov}}{d\text{Nasdaq}} + \frac{\partial\text{Metric}}{\partial \text{NormInst}} \frac{d\text{NormInst}}{d\text{Nasdaq}} \\
+ \frac{\partial\text{Metric}}{\partial \text{logMCap}} \frac{d\text{logMCap}}{d\text{Nasdaq}} \frac{d\text{logShrsOut}}{d\text{Nasdaq}}
\]

\[
= \gamma_{\text{Nasdaq}} \\
+ \gamma_{\text{KOLiq}} \alpha_{\text{Nasdaq}} + \gamma_{\text{NormSI}} \theta_{\text{Nasdaq}} \\
+ \gamma_{\text{AnCov}} \zeta_{\text{Nasdaq}} + \gamma_{\text{NormInst}} \psi_{\text{Nasdaq}} \\
+ \gamma_{\text{logMCap}} \xi_{\text{Nasdaq}} + \gamma_{\text{logShrsOut}} \omega_{\text{Nasdaq}}
\]
\[
\frac{d\text{Metric}}{d\text{Disp}} = \frac{\partial\text{Metric}}{\partial\text{Disp}} + \frac{\partial\text{Metric}}{\partial\text{KOLiq}} \frac{d\text{KOLiq}}{d\text{Disp}} + \frac{\partial\text{Metric}}{\partial\text{NormSI}} \frac{d\text{NormSI}}{d\text{Disp}} + \frac{\partial\text{Metric}}{\partial\text{AnCov}} \frac{d\text{AnCov}}{d\text{Disp}} + \frac{\partial\text{Metric}}{\partial\text{NormInst}} \frac{d\text{NormInst}}{d\text{Disp}} + \frac{\partial\text{Metric}}{\partial\text{logMCap}} \frac{d\text{logMCap}}{d\text{Disp}} + \frac{\partial\text{Metric}}{\partial\text{logShrsOut}} \frac{d\text{logShrsOut}}{d\text{Disp}}
\]

\[
= \gamma_{\text{Disp}} + \gamma_{\text{KOLiq}} \alpha_{\text{Disp}} + \gamma_{\text{NormSI}} \theta_{\text{Disp}} + \gamma_{\text{AnCov}} \zeta_{\text{Disp}} + \gamma_{\text{NormInst}} \psi_{\text{Disp}} + \gamma_{\text{logMCap}} \xi_{\text{Disp}} + \gamma_{\text{logShrsOut}} \omega_{\text{Disp}}
\]

\[
\frac{d\text{Metric}}{d\text{SSCC}} = \frac{\partial\text{Metric}}{\partial\text{SSCC}} + \frac{\partial\text{Metric}}{\partial\text{KOLiq}} \frac{d\text{KOLiq}}{d\text{SSCC}} + \frac{\partial\text{Metric}}{\partial\text{NormSI}} \frac{d\text{NormSI}}{d\text{SSCC}} + \frac{\partial\text{Metric}}{\partial\text{AnCov}} \frac{d\text{AnCov}}{d\text{SSCC}} + \frac{\partial\text{Metric}}{\partial\text{NormInst}} \frac{d\text{NormInst}}{d\text{SSCC}} + \frac{\partial\text{Metric}}{\partial\text{logMCap}} \frac{d\text{logMCap}}{d\text{SSCC}} + \frac{\partial\text{Metric}}{\partial\text{logShrsOut}} \frac{d\text{logShrsOut}}{d\text{SSCC}}
\]

\[
= \gamma_{\text{SSCC}} + \gamma_{\text{KOLiq}} \alpha_{\text{SSCC}} + \gamma_{\text{NormSI}} \theta_{\text{SSCC}} + \gamma_{\text{AnCov}} \zeta_{\text{SSCC}} + \gamma_{\text{NormInst}} \psi_{\text{SSCC}} + \gamma_{\text{logMCap}} \xi_{\text{SSCC}} + \gamma_{\text{logShrsOut}} \omega_{\text{SSCC}}
\]
\[
\frac{\text{dMetric}}{\text{dTCC}} = \frac{\partial \text{Metric}}{\partial \text{TCC}} + \frac{\partial \text{Metric}}{\partial \text{OLiq}} \frac{\text{dOLiq}}{\text{TCC}} + \frac{\partial \text{Metric}}{\partial \text{NormSI}} \frac{\text{dNormSI}}{\text{TCC}} + \frac{\partial \text{Metric}}{\partial \text{AnCov}} \frac{\text{dAnCov}}{\text{TCC}} + \frac{\partial \text{Metric}}{\partial \text{NormInst}} \frac{\text{dNormInst}}{\text{TCC}} + \frac{\partial \text{Metric}}{\partial \logMCap} \frac{\text{dlogMCap}}{\text{TCC}} + \frac{\partial \text{Metric}}{\partial \logShrsOut} \frac{\text{dlogShrsOut}}{\text{TCC}}
\]

\[
= \gamma_{\text{TCC}} + \gamma_{\text{OLiq}} \theta_{\text{TCC}} + \gamma_{\text{NormSI}} \theta_{\text{TCC}} + \gamma_{\text{AnCov}} \xi_{\text{TCC}} + \gamma_{\text{NormInst}} \psi_{\text{TCC}} + \gamma_{\logMCap} \xi_{\text{TCC}} + \gamma_{\logShrsOut} \omega_{\text{TCC}}
\]

I perform each of my analyses for 2014-September 2018 and find, as shown in Table 3 below,\textsuperscript{27} that FINRA 2241’s impact on each of the above ten systematic and objective market quality metrics was economically insignificant. If one were to redefine economic significance as the coefficient being greater than 0.001 in absolute value, or even greater than 0.0001 in absolute value, even then, the impact of FINRA 2241 would still be economically insignificant.

\textsuperscript{27}As explained in Subsection 3.9, the total derivatives are all comparable, \(\frac{\text{d}}{\text{dT}} = \beta(y,x)\) means that a 1% increase in the cumulative probability of \(x\) causes a \(\beta(y,x)\) % increase in the cumulative probability of \(y\), and similarly, \(\frac{\text{d}}{\text{dI}} = \beta(y,1)\) where \(I\) is an indicator variable means that there is a \(\beta(y,1)\) higher cumulative probability of \(y\).

I indicate economically significant positive impacts \((\frac{\text{d}}{\text{dT}} = \beta(y,x) > 0.01)\) by green highlighting and economically significant negative impacts \((\frac{\text{d}}{\text{dT}} = \beta(y,x) < -0.01)\) by red highlighting.
5 Conclusions

I calculated six measures of market efficiency, based on intraday event studies, and also measured diligence, objectivity, quality, and accuracy, of forecasts by an analyst-analyst-firm. I used the panel nature of the data to identify appropriate instruments for the endogenous variables and for the variables that are measured with error, and I used Three Stage Least Squares (3SLS) and Errors in Variables (EiV) to estimate the seven-equation structural model developed by Bhattacharya (2022). I developed the methodology to calculate the total impact of a factor which is the sum of the direct impact and the indirect impacts, and by using the Gaussian cumulative distribution of the $Z$-score of each variable (except for an indicator variable), I made the impacts comparable. I found that FINRA 2241’s impact on each of the above ten objective and systematic market quality metrics was economically insignificant.

References

Aman, H., Beekes, W., Berkman, H., Bohmann, M., Bradbury, M., Chapple, L., Chang, M., Smith, T., et. al. (2019). Responsible Science: Celebrating the 50-Year Legacy of Ball and Brown (1968) Using a Registration-Based Framework. Pacific-Basin Finance Journal, 56(C), 129-150


Bhattacharya, R. & Gupta, M. (2022). Diligence, Objectivity, Quality, and Accuracy. SSRN


Law, K. (2021). Goodbye to IBES (or Not?). *SSRN*


Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table1.xlsx
- Table2.1.xlsx
- Table2.2.xlsx
- Table3.xlsx
- Data.docx