IIMB-WP No. 607/2020

WORKING PAPER NO: 607

The Pricing of Earnings in the Presence of Informed Trades: A Simple GMM Approach

Murugappa (Murgie) Krishnan
Radford University, 801 E Main St,
Radford, VA 24141, United States,
murgie@gmail.com

Srinivasan Rangan
Associate Professor
Finance & Accounting
Indian Institute of Management Bangalore
Bannerghatta Road, Bangalore – 5600 76
srinivasanr@iimb.ac.in

Nikhil Vidhani
Doctoral Student
Finance & Accounting
Indian Institute of Management Bangalore
Bannerghatta Road, Bangalore – 5600 76
nikhil.vidhani16@iimb.ac.in

Year of Publication – March 2020

The authors acknowledge the support of the NSE-NYU Stern School of Business Initiative for the Study of the Indian Capital Markets. The views expressed in this working paper are those of the authors and do not necessarily represent those of NSE or NYU. We sincerely thank Nagpurnanand Prabhala (discussant), Venky Panchapagesan, Arnab Bhattacharya (discussant), Sudershan Kuntluru (discussant), Jordi Caballé, Pete Kyle, Bharat Sarath, Vish Viswanathan, and seminar participants at the 2015 NSE-NYU Conference, IIT-Madras, IIT-Kharagpur, IIM Bangalore, CAFRAL at RBI, the IIM-Calcutta Finance Research Workshop, and the third JAAF-India Conference, for their comments and suggestions. We also acknowledge the help of Shyam Benegal, whose parliamentary question first led to the release by SEBI of Foreign Institutional Investor trading data in India.

IIMB-WP No. 607/2020

The Pricing of Earnings in the Presence of Informed Trades: A Simple GMM Approach

Abstract

We build a Kyle-type pricing model with earnings and trading signals and estimate its deep parameters - the information advantages of traders and firms, the correlation between the firm and traders' information, and the noise variance. Moment conditions derived from the pricing rule yield a simpler form than in prior work, and we validate our model both asymptotically and in a finite sample. For our sample from Indian markets, we find that traders know more about firm payoffs than firms themselves. For many firms, the market's weight on unexpected earnings is negative, causing good news to be bad news.

Keywords: Foreign Institutional Investors, GMM, Institutional Trading, Kyle Model, Earnings Announcements.

1. Introduction

Investors rely on multiple signals about a firm when setting its prices. Firm-initiated *public* announcements of earnings are one class of signals that have been shown to be priced. In addition, investors also learn from the trades of informed traders. Presumably, these traders possess *private* information that is correlated with firms' future payoffs. While the price-informativeness of each of the two signals (firm news and informed trades) is well-accepted, little is known about how investors price them when they are simultaneously present. The central contribution of this paper is to develop and empirically test a model of the pricing of firm news *and* informed trades when (a) both firms and traders have information about firm prospects and (b) their information is potentially correlated.

The point of departure in the paper is a single-period model of asset pricing under imperfect competition. The model features four types of players: the firm that reports earnings, a strategic trader, noise traders, and competitive market makers. Price is a linear function of the two public signals – unexpected earnings and the trader's order flow. Both the firm and the trader have information about the firm's payoffs and these payoffs are the sum of component innovations, a la Admati and Pfleiderer (1988), one component known to the firm, and another, to the strategic trader. The components can be interpreted as the information advantage that the firm (trader) has relative to what the market has based on priors alone.[1] Allowing these components to be correlated lets us capture, within a parsimonious framework, a variety of relationships between the firm and the trader's information.

The empirical literature relating to Kyle's λ and the PIN measure (the probability of informed trading) derived from the Glosten-Milgrom model has interpreted the estimated parameters assuming that the underlying model holds. In contrast, we validate our model both asymptotically with the Hansen-Sargan J-statistic and in a finite sample. Our innovation is to use moment conditions derived from the equilibrium pricing rule. These allow for a very simple GMM strategy. Nonlinearity in parameters, and lack of priors about parameters relating to private information that have not been estimated before, pose a challenge. Important in enabling model validation is obtaining better starting points. We do so with an initial evaluation of the objective function over a dense grid. In the estimation we use a perturbation modification of the standard GMM algorithm which reduces the risk of a gradient-search algorithm being trapped in a poor local solution.

[1] For example, for firms, it is reasonable to assume that their information advantage would be about the earnings that they report, and in particular, about the new information in earnings. For strategic traders, such as institutions, the information advantage is likely to arise from their superior ability in interpreting information outside the firm. Examples of such information include macroeconomic data, government policy, industry factors, and even the state of financial markets.

We show that how unexpected earnings affects price depends on the correlation between the firm and strategic trader's information and the information advantage of the trader relative to the firm. Thus, compared to previous work, our model predicts that the market values unexpected earnings not just because of payoff information contained in earnings (the direct effect), but also because of what earnings tells the market about what traders know about firm payoffs (the indirect effect). A second result, which is fairly intuitive, is that the pricing of order flow is increasing in the informedness of the trader but decreasing in the noisiness of her trades. Overall, we obtain a model of price that is linear in unexpected earnings and the strategic trader's order flow and a nonlinear function of three unobservable (primitive) parameters – the correlation between firm's information about earnings and the strategic trader's private information, the relative information advantage of the strategic trader, and the variance of the level of noise trading.

We implement our tests for a sample of earnings announcements in India for the years 2003-2016 (over 1,100 firms and 17,800 announcements). Our model requires proxies for price, earnings news, and order flows. We measure price impact as the two-day abnormal return over the earnings announcement period and define earnings news as the difference between quarterly earnings per share and its four-quarter lagged value. To measure order flow, we employ net buying by Foreign Institutional Investors (FIIs) over the earnings announcement period. Our motivation to study FII trades stems from three reasons. First, FIIs are a significant player in Indian markets, accounting for twenty to thirty percent of total turnover on the two leading exchanges (National Stock Exchange (NSE) and Bombay Stock Exchange (BSE)). Second, the availability of a database on daily FII-level trades that is public and free allows for easy replication. Third, prior evidence on the informedness of FIIs has examined the correlation between their trades and subsequent returns. In contrast, we estimate a pricing model from which we uncover a parameter that speaks to their contemporaneous information advantage relative to price-setters with only public information. Thus, we contribute to the debate on the informedness of FIIs.

Our results indicate that, on average, the information advantage of FIIs with respect to the component they have information about exceeds the information advantage that firms have with respect to information released via earnings announcements. We also find that correlation (ρ) between the two fundamental information components is generally positive. This establishes that the response to earnings is in part the response to anticipated private information of traders. Learning from prices or order flows about agents' private information has been studied a lot. That other public signals, like earnings, may also reveal traders' private information in a correlated environment, has not been sufficiently recognized so far. Additionally, the information advantage of the FIIs is larger than the noise in their trades. In contrast to most papers studying institutional trades, we test and adjust for the endogeneity of these trades. Further, our conclusions are robust to the inclusion of earnings announcements with no FII trades.

To assess if there is heterogeneity in primitive parameter estimates, we estimate the pricing model for a sub-sample of 365 firms with at least twenty time-series observations. Our results indicate that firms display considerable variation in the deep parameters. Interestingly, while the correlation parameter is positive on average, it is negative for a subset of firms. For this subset, the combination of the negative correlation with the relatively higher information advantage of the FIIs, causes good news about firm earnings to be viewed as bad news by markets, as noted in a different setting by Lundholm (1988) and Manzano (1999). The traditional result that unexpected earnings are valued positively by markets may reflect the omission of a key market signal, institutional trades. This is more than a theoretical curiosity since the negative valuation of good news obtains for 102 out of the 365 firms (28%). This conclusion is possible only because we explicitly model the underlying equilibrium in a correlated environment and confront that model with data.

We also examine whether primitive parameters vary with certain firm and trader characteristics. Estimates of the traders' information advantage and market noise, rather than just λ, their ratio in traditional Kyle models, and adjusting for the endogeneity of trades, yields new insights. We find that the information advantage of FIIs is increasing in firm size suggesting that traders pay more attention to larger firms. The FIIs' information advantage is lower for loss firms compared to profitable firms, consistent with traders having more difficulty learning about loss firms. Firms with small profits have a negative correlation between what firms and traders know. This is evidence of more disagreement between them in interpreting common information. Lastly, we examine the effect of trader attention on FIIs' information advantage by partitioning our sample based on the average number of market-wide earnings announcements over the earnings announcement period. Following Hirshleifer, Lim, and Teoh (2009), we assume that investor attention decreases when the number of market-wide announcements increases. The results indicate that, as one would expect, both the absolute and relative information advantage of the FIIs increases as their attention increases.

In additional analyses, we consider the impact of non-announcement period information on our results. By comparing our estimates in the main model with benchmark models that have only an earnings signal, or only an FII trading signal, we find that earnings and FII trading are mutual information complements. FII trading is a complement to earnings because its presence causes the market to use earnings also to learn about FIIs' private information. Earnings are a complement to FII trading because with earnings the reduction in market noise is more than the reduction in the FIIs' information advantage.

One contribution we make is to quantify various aspects of trader behavior that so far have only been discussed speculatively. It is not easy to have priors about parameters like the traders' information advantage and the correlation between firm and traders' information. Our work provides empirical measures of such parameters, and so can open the way to additional questions being addressed. We also contribute to the literature that examines whether institutional trading predicts the sign of subsequent earnings news or returns by providing an alternative way to characterize smart institutions – the size of the relative information advantage of institutions at the time of earnings announcements.[2]

The rest of the paper proceeds as follows. We review related literature, develop the pricing model, and describes its equilibrium properties, discuss design considerations to estimate the model, provide variable definitions, describe data sources, present results, and then conclude.

2. Prior Literature

2.1. Theoretical Work

Admati (1985) generalized the single-security noisy rational expectations model under perfect competition due to Hellwig (1980) to the case of multiple securities, allowing for general variance-covariance matrices governing payoffs, errors in private signals, and liquidity noise. She showed that a common intuition in a single-security setting that a security price would be increasing in its own payoff need not hold with many securities and sufficient correlation. Caballé and Krishnan (1994) generalized the risk-neutral imperfect competition model due to Kyle (1985), to the case of N assets and K traders, with a similarly rich correlation structure, and showed that asset prices again need not be increasing in their own payoffs. They also show that in a correlated environment, portfolio diversification can arise for a reason unrelated to risk: to minimize the revelation of information.

[2] The conclusion from most studies in this literature is that institutional investors are informed, and profit from their trades; their net buying is positively associated with subsequent stock returns (e.g., Nofsinger and Sias (1999); Gompers and Metrick (2001); Yan and Zhang (2009); Campbell, Ramadorai, and Schwartz (2009); Puckett and Yan (2011); and Hendershott, Livdan, and Schurhoff (2015)). But there are also papers with contrary evidence (e.g., Bushee and Goodman (2007); Griffin, Shu, and Topaloglu (2012); and Edelen, Ince, and Kladlec (2016)).

Lundholm (1988), under perfect competition, and Manzano (1999), under imperfect competition, show that a similar ambiguity can arise even with one security if there were multiple signals available; for example, a public signal like earnings together with private signals for each trader. A security's price may not increase in earnings. The key to this result is the information structure used in both these papers, where an asset has a payoff v and the signals, public and private, are of the form $s_i = v + e_i$, with Cov(e_i, e_j) = C, $i \neq j$, C not necessarily zero. In this case, each signal has both a direct and an indirect effect. A large value of s_i could indicate a high v, and this is the direct effect. On the other hand, it could indicate a large e_i, and, if the covariance between errors in signals is high enough, also a high $e_j, j \neq i,$ and consequently a lower v; this is the indirect effect. When the indirect effect dominates, a large value of s_i (good news) can be bad news. We obtain a similar ambiguity in the sign of the coefficient on the public signal in our model, but it arises from a combination of a negative correlation between payoff components and a greater information advantage for traders.

Davila and Parlatore (2018), in a spirit similar to this paper, consider the estimation of measures of price informativeness within a linear-demand framework. In one of their examples, they impose more structure and solve for some primitive parameters. The difference between their work and ours is that our model has more parametric structure. Further, we estimate all of the model's primitive parameters: the precision of traders' private information, its correlation with the firm's information, and the variance of the background market noise.

2.2. Empirical Work

Our paper is related to previous empirical work on the investment performance of FIIs in several countries, including Finland, Indonesia, Japan, South Korea, and Taiwan). While Grinblatt and Keloharju (2000) and Huang and Shiu (2009) conclude that FIIs generate superior performance, Kang and Stulz (1997), Dvořák (2005), and Choe, Kho, and Stulz (2005) report the opposite. In India, while there are many news stories and anecdotes of FIIs' importance, formal evidence is scarce. Acharya, Anshuman, and Kumar (2014) find that stocks with high FII order flow innovations experience a coincident price increase that is permanent, whereas stocks with low innovations exhibit a coincident price decline that is in part transient, reversing itself within two weeks. The results are consistent with price pressure on stock returns induced by FII sales, as well as information being revealed, as in our model, through FII purchases and sales.

Our paper is also related to the few studies that examine study prices, earnings, and actual daily institutional trades jointly. Daley, Hughes, and Rayburn (1995) study the effect of block trades during earnings announcements and ask if anticipated public announcements give rise to private information acquisition, and permanent price effects. Campbell, Ramadorai, and Schwartz (2009) show that inferred institutional trades anticipate earnings surprises and the post-earnings announcement drift. Hu, Ke, and Yu (2018) report that transient institutions sell in response to small negative surprises at earnings announcements which in turn improves the informational efficiency of share prices. We formulate and empirically test a model of the pricing of firm earnings *and* informed trades when (a) both firms and traders have information about firm prospects and (b) their information is potentially correlated.

3. Model

We formulate a single-period model of asset pricing under imperfect competition with both public and private signals. The model is based on Kyle (1985) and Rochet and Vila (1994). Our objective is to create a model that is sufficiently rich and yet yields a simple pricing rule with testable implications.

3.1. Assumptions

(A1) Assets, asset payoffs, and information about asset payoffs:

There is one risky asset and one riskless asset (numeraire) with a payoff and price equal to one. The payoff to the risky asset (and equivalently, information about this payoff) is given by $\tilde{v}$ and is expressed as the sum of two informational innovation components: $\tilde{v} = \tilde{v}_F + \tilde{v}_T$.[3] Here, $\tilde{v}_F$ is the component for which the firm has an information advantage relative to others, and $\tilde{v}_T$ is the component for which a strategic trader has an information advantage. We assume that the components $\tilde{v}_i \sim N(0, \sigma_i^2), i = F, T$, with $\text{cov}(\tilde{v}_F, \tilde{v}_T) = \rho . \sigma_F . \sigma_T, \rho \in (-1,1), \sigma_i > 0, i = F, T$.

The component structure for payoffs has been used before in other papers, including Admati and Pfleiderer (1988). Unlike Admati and Pfleiderer (1988), we allow the components to be correlated. Thus, the components can be substitutes, complements, or independent of each other. In sum, we employ a more general structure than prior research did. Yet, our model is parsimonious in terms of the number of parameters to be estimated.

[3] The label information innovation component is deliberately used to highlight the idea that our variables $\tilde{v}_F$ and $\tilde{v}_T$ need not represent cash flow payoffs but are signals about such payoffs.

How should one interpret $\tilde{v}_i$? We could interpret $\tilde{v}_i$ as perfect information on a component observable to agent $i, i = F, T$, and for ease of exposition, we will sometimes do that. But it will be more convenient to interpret $\tilde{v}_i$ as the posterior information advantage of agent i relative to prior beliefs, i.e., $\tilde{v}_i = E(\tilde{v} \mid \tilde{I}_i) - E(\tilde{v})$, where I_i is the information set of agent i.[4] In our empirical work, we do not specify the information sets, I_i. Instead, we estimate the variances of the distributions of $\tilde{v}_F$ and $\tilde{v}_T$. Thus, we avoid having to estimate additional parameters related to the players' information sets while being slightly more general.

For firms, it is reasonable to assume that their information advantage would be about the earnings that they report, and in particular, about the new information in earnings. Consistent with this idea, in our empirical work, we equate σ_F^2, the information advantage of the firm, to the variance of unexpected earnings, which can be estimated directly from the data. For strategic traders, such as institutions, the information advantage is likely to arise from their superior ability in interpreting information outside the firm. Examples of such information include macroeconomic data, government policy, industry factors, and even the state of financial markets. But this interpretation is only a suggestion. It is not essential. Our empirical estimation uncovers the information advantage of strategic traders, σ_T^2. By comparing estimates of σ_F^2 and σ_T^2, we provide a simple way of describing whether firms know more or less than do traders.

(A2) Agents:

Firm: There is a firm, denoted by subscript $i = F$, which observes $\tilde{v}_F = v_F$ perfectly and reports it faithfully, as required to do so under accounting rules. Note, however, that because of the component structure of total firm payoff, seeing and reporting perfect information on one component is not the same as knowing and reporting "everything." Our assumption A1 above allows firms to know a lot or little. One interpretation is that auditing works and results in compliance (see, e.g., Shin (1994)). Alternatively, we could invoke models of cheap talk (for example, Bhattacharya and Krishnan (1999)) in which firms have an incentive to make truthful disclosures, despite being able to lie with impunity. In either case, this assumption is broadly consistent with the vast literature that has documented a consistent positive association between unexpected earnings and abnormal returns, while also noting that only a small portion of price variation is explained by earnings variation, even within an earnings announcement window.

---

[4] For a convenient summary of the algebra of informational advantages, see the remarks following assumption (A3) in Caballé and Krishnan (1994).

Noise trader: This trader generates a random net demand of $\tilde{z}$, $\tilde{z} \sim N(0, \sigma_z^2)$, $\tilde{z}$ is uncorrelated with the two payoff components. This assumption enables us to capture the non-strategic or non-information-based activity of traders.[5]

Strategic trader: This trader, denoted by subscript $i = T$, is strategic and informed. She chooses a demand for the risky security, x, based on all information available to her: the public signal created by the firm's earnings announcement, v_F, the perfect private signal about the second component, $\tilde{v}_T = v_T$, and the noise trades of some other traders, z. Being able to observe z is different from Kyle (1985), but similar to Rochet and Vila (1994). Therefore, the strategic trader is not just better informed than the market makers (who can only observe the aggregate demand, $\omega = x + z$), but their information is nested in hers.[6]

Competitive market makers: We assume that there are competitive risk-neutral market makers whose competition makes them earn zero expected profits. Hence, the price they set for the risky asset is equal to the expected payoff from the security given all publicly available information. We assume that public information will consist of the earnings signal that the firm provides, v_F, and the aggregate order flow, $\omega = x + z$. Hence, the price $p = E(\tilde{v} \mid v_F, \omega)$. We also assume a linear pricing rule, $p = \alpha + \beta v_F + \lambda \omega$. Given the uniqueness of equilibrium result in Rochet and Vila (1994), this ex-ante assumption of a linear pricing rule is not an additional restriction but makes the solution procedure more convenient.

3.2. Definition of equilibrium

An equilibrium of this model is defined by a trader strategy $x(v_F, v_T, z)$ and a pricing rule $p = \alpha + \beta v_F + \lambda \omega$, such that we have

(i) Trader optimization: Given the above pricing rule, and any triple of realized values $\{v_F, v_T, z\}$ the trader T has a demand strategy $x(v_F, v_T, z)$ that is at least as good as any alternate strategy $x'(v_F, v_T, z)$.

(ii) Market efficiency: for any realization of earnings v_F and aggregate order flow $\omega = x + z$, the price $p = E(v|v_F, \omega)$.

[5] We have studied a variant of the model that allows for z to be correlated with a payoff component. It is possible to compute equilibrium even in such a model, but the added analytical complexity yields no additional intuition about trader or market behavior and complicates parameter estimation substantially.

[6] Rochet and Vila (1994) adopt this assumption for an important theoretical reason. Given a nested information structure, and exogenous total profits in the game, they show uniqueness of equilibrium under otherwise very general assumptions. In the Kyle (1985) framework, uniqueness has only been shown given a linear pricing rule, and uniqueness of equilibrium in general is still an open question.

3.3. Equilibrium in benchmark models

We first define two simple benchmark models that help interpret the results from our main model. In the first model (Regime 1) there is no strategic trading, and the only signal available to market makers is v_F. It is evident given our other assumptions that the following holds under Regime 1:

Lemma 1: The equilibrium price $p = v_F$, so $\beta = 1$.

In the second benchmark model (Regime 2), we have strategic trading but no earnings announcements. That is, price is set based on strategic trading in non-announcement periods. Given our other assumptions, this model closely resembles an example in Rochet and Vila (1994),[7] but for the component payoff structure. It is straightforward to show the following:

Lemma 2: The unique equilibrium of this model is defined by a trader strategy $x(v_T, z)$ $= \tau_0 + \tau_1 v_T + \tau_2 z$, where $\tau_0 = 0,\ \tau_1 = \left(\frac{\sigma_z}{2\sigma_T}\right),\ \tau_2 = -\left(\frac{1}{2}\right)$, and a pricing rule $p = \alpha + \lambda\omega$, where $\alpha = 0,\ \lambda = \left(\frac{\sigma_T}{\sigma_z}\right)$.

3.4. Properties of equilibrium

In this sub-section, we focus on the main model with both earnings and trading signals (the Regime 3 model).

Proposition 1: For $\rho \in (-1,1)$ the unique equilibrium of this model is defined by a trader strategy $x(v_F, v_T, z) = \tau_0 + \tau_1 v_F + \tau_2 v_T + \tau_3 z$, where $\tau_0 = 0,\ \tau_1 = \left(\frac{-\rho\ \ z}{2\sigma_F(\sqrt{1-\rho^2})}\right),\ \tau_2 = \left(\frac{\sigma_z}{2\sigma_T(\sqrt{1-\rho^2})}\right),\ \tau_3 = -\left(\frac{1}{2}\right)$, and a pricing rule $p = \alpha + \beta v_F + \lambda\omega$, where $\alpha = 0,\ \beta = 1 + \rho\left(\frac{\sigma_T}{\sigma_F}\right)$, $\lambda = \left(\frac{\sigma_T(\sqrt{1-\rho^2})}{\sigma_z}\right)$.

The proof is outlined in Appendix A. In the key final step, we equate coefficients in the pricing rule, to get three equations of the form, $\alpha = f_1(\alpha, \beta, \lambda)$, $\beta = f_2(\alpha, \beta, \lambda)$, $\lambda = f_3(\alpha, \beta, \lambda)$. From the first alone, it is easy to show that $\alpha = 0$. Manipulating the other two leads to a cubic in two variables - β and λ, instead of in λ alone, as in Kyle (1985) and Rochet and Vila (1994). Of the three solutions, only

[7] In a Kyle (1985) setup, the expression for λ would have a coefficient of (1/2).

one satisfies $\lambda > 0$, which is needed to satisfy the second-order conditions. Therefore, we have a unique real root. The solution is easily verified.

In Proposition 1, the intercepts (τ_0 and α) being zero reflects the zero-mean priors for all variables. In the strategic trader's demand function, the coefficient τ_1, which is the weight that the trader places on v_F, is influenced by the correlation between the two payoff components. Though v_F provides perfect information about one component and is public, the expression for τ_1 is complex because v_F also yields information about the second component, as $E(v_T|v_F) = \rho\left(\frac{\sigma_T}{\sigma_F}\right)v_F$. The coefficient τ_1 is increasing in the ratio $\left(\frac{\sigma_z}{\sigma_F}\right)$ for any $\rho < 0$; decreasing in that ratio, for any $\rho > 0$, and is decreasing in ρ. The coefficient τ_2 (the weight the strategic trader places on v_T) is increasing in the ratio $\left(\frac{\sigma_z}{\sigma_T}\right)$ for any ρ. As noise (σ_z) increases, the greater camouflage encourages the trader to be more aggressive. Also, as her information advantage (σ_T) increases, the market will place more weight on the order flow, inducing the trader to reveal less by becoming less aggressive.

In the pricing rule, β, the weight on unexpected earnings (v_F), is the earnings response coefficient. The expression for β includes the coefficient of v_F in the conditional expectation, $E(v_T|v_F)$, that is, $\rho\left(\frac{\sigma_T}{\sigma_F}\right)$, besides 1, the coefficient when only v_F is available as a signal (see Lemma 1 above). Thus, the expression tells us that the earnings response coefficient depends not only on what a firm reveals about v_F but also on what the market learns from the firm's report about what traders know (v_T).

The weight on the aggregate order flow (ω) is denoted by λ. That we need $\lambda > 0$ follows from the second-order condition. If this did not hold, by buying more a trader would push the price not up but down, till she would want to hold an arbitrarily large position paying nothing. Relative to the benchmark case without v_F (see Lemma 2 above), the expression for λ reflects the presence of that second possibly correlated signal. When $\rho \neq 0$, observing v_T confers less of an information advantage to the traders, relative to market makers, who can now guess part of the strategic trader's information. The market makers, therefore, set a flatter pricing rule than they would if the information asymmetry is greater. In the limit, as all of the trader's information is anticipated, she has no information advantage.

Our model has implications for the correlation between the two public signals. Given v_F, the trader effectively faces both a different intercept and slope, and her demand reflects her information advantage, defined by the residual $v_T - E(v_T|v_F)$, which is orthogonal to v_F. Since the order flow is only a noisy linear transformation of this orthogonal residual, in equilibrium, v_F and order flow ω are also orthogonal. This orthogonality obtains, even though the correlation parameter ρ could be non-zero.

The correlation parameter features both in our model and that of Lundholm (1988). In Lundholm (1988), the linear pricing rule involves private signals and earnings news. Because of the additional observable, order flows, in the pricing rule, we have an easier estimation problem. Also, because we assume risk neutrality, our equilibrium expressions are simpler than those in Lundholm (1988) and Manzano (1999), who assume risk aversion; and so estimation of primitive parameters becomes easier.

Can β be negative? From the expression for β, $\beta < 0 \Leftrightarrow \rho\left(\frac{\sigma_T}{\sigma_F}\right) < -1$. This can arise only if we have (i) $\rho < 0$, and (ii) for negative ρ, we also have $\sigma_T > \sigma_F$ by a sufficient margin. To interpret this, consider an equivalent setting where the total payoff $\tilde{v} = \tilde{v}_1 + \tilde{v}_2 + \tilde{v}_C$, $\tilde{v}_F = \tilde{v}_1 + a * \tilde{v}_C$, $\tilde{v}_T = \tilde{v}_2 + (1-a) * \tilde{v}_C$. The correlation arises because of the common component $\tilde{v}_C$ and will be negative when $a < 0$ or $a > 1$. For $\beta < 0$, besides the firm and the strategic trader having a common component about which they disagree, it must also be the case that the trader's informational advantage (σ_T) must be sufficiently larger than that of the firm (σ_F). A practical implication of this for empirical work is that an estimate of the shallow parameter $\beta < 1$ immediately tells us that traders must know more than firms.[8]

4. Variable Measurement, Identification, and Estimation Method

Our interest in this paper is in estimating and analyzing the parameters of the linear pricing rule obtained under Proposition 1 (the Regime 3 model):

$$p = \alpha + \beta v_F + \lambda\omega, \text{ where}$$
$$\alpha = 0,\ \beta = 1 + \rho\left(\frac{\sigma_T}{\sigma_F}\right),\ \lambda = \left(\frac{\sigma_T(\sqrt{1-\rho^2})}{\sigma_z}\right).$$

We estimate this pricing model using multiple methods – OLS, 2SLS, and GMM. Note that the model is nonlinear in its parameters (though still linear in variables). Hence, the linear estimation methods - OLS and 2SLS, will allow estimation of the shallow parameters, β and λ, but not the deep parameters.

[8] The reason for a possible counter-intuitive sign (good news is interpreted as bad news) is different in this setting from the reason in Lundholm (1988) and Manzano (1999). In those papers, the multiple signals are all about the same component and the correlation governs the error covariance, so β (coefficient of the public signal in the price function) in those models can be negative when the indirect effect dominates the direct effect for sufficiently large positive error covariance.

GMM helps estimate the deep parameters and also makes it convenient to adjust for the endogeneity of order flows.

4.1 Variable Measurement

We measure the price impact p as the abnormal return over the earnings announcement period (ERET), v_F as unexpected earnings (UE), and ω, the order flow, as net buying by Foreign Institutional Investors (FIIs) over the earnings announcement period (FIITR).[9] Thus, our empirical specification for firm i and quarter t is:

$$ERET_{it} = \alpha + \beta \times UE_{it} + \lambda \times FIITR_{it} + \varepsilon_{it} \qquad (1)$$

where $\alpha = 0, \beta = 1 + \rho\left(\frac{\sigma_T}{\sigma_F}\right), \lambda = \left(\frac{\sigma_T(\sqrt{1-\rho^2})}{\sigma_z}\right)$,

ERET = Abnormal Return compounded over the day of the earnings announcement and the following day, (0,1),

UE = Earnings per Share in quarter t less Earnings per Share from four quarters prior, scaled by share price on the last date of the quarter for which earnings is announced, and

FIITR = Net Buying by *all* FIIs over days [0, 1] divided by shares outstanding. Net FII buying for a firm on a day equals the number of shares bought less the number of shares sold for that firm by all FIIs on that day.

To measure the dependent variable ERET, we obtain the earnings announcement date (day 0) and returns on day 0 and 1. We treat the date of the board meeting on which financial results are approved as day 0.[10] ERET is defined as the residual from a panel regression of the two-day raw return during the earnings announcement period on twelve control variables, firm fixed effects, and year effects.[11] Our first set of control variables is drawn from prior work on asset pricing. Fama and French (2016) show that five factors explain a significant fraction of the cross-section of monthly returns. The factors are market-wide return, firm size, book-to-market ratio, operating profitability scaled by assets, and prior asset growth. We conjecture that these factors would explain the cross-section of earnings announcement returns as well. Our second set of seven controls are other firm characteristics that have

[9] We model returns, rather than price, because our model is about the impact of new information.

[10] As per BSE and NSE Regulations, listed firms are required to file their quarterly financial results within thirty minutes of the end of the board meeting, presumably to reduce the likelihood of illegal insider trading. See item 13 in the URL, https://beta.bseindia.com/corporates/compliancecalendar.aspx.

[11] An alternate specification, where we include the twelve control variables as additional regressors in a model of the two-day compounded raw return, does not change any of our conclusions. That is, the parameter estimates when we estimate a model of abnormal return on UE and FIITR are very similar to those from a model of raw return on UE, FIITR, and the twelve control variables.

been shown to be related to institutional trading (Gompers and Metrick (2001); Yan and Zhang (2009)). Whether these characteristics are related to earnings announcement returns is an open question. However, we include them as regressors for a pragmatic reason - to reduce the likelihood of any omitted variable bias. The seven additional controls are three-month prior return, prior return volatility, prior monthly volume, lagged UE, firm age, lagged annual dividend yield, and beginning of quarter price. Detailed variable definitions are contained in Appendix B.

Our measure of unexpected earnings per share, UE, assumes that the market uses earnings from four quarters before as its expectation when pricing the firm (Bernard and Thomas (1990)). As an alternative, we also compute unexpected earnings as the difference between earnings per share and the mean analysts' forecast of earnings per share before the earnings announcement, scaled by the quarter-end price (AFE). Unfortunately, analysts' forecast data are available only for a sub-sample of firm-quarters in India. Hence, in our conclusions, we emphasize the findings based on the larger sample using UE.

We measure order flow as net buying by all foreign institutional investors (FIIs) over the earnings announcement period (FIITR). Because we examine only FII trades, we implicitly regard trading by non-FIIs as being of at best second-order importance. This choice is driven by data availability; unfortunately, we do not have data on daily non-FII trades.[12]

Order flows, $\tilde{\omega}$, are endogenous in the theory; therefore, the corresponding data variable FIITR is endogenous. We report 2SLS estimates of the pricing rule to address this endogeneity. In our 2SLS estimation, we use two instruments for FIITR - the change in the US Dollar-Rupee rate over the week ending on day -1 relative to the earnings announcement period (CH_WK_EXCH) and the market return over the week ending on day -1 (WK_MRET).

We also estimate an empirical model where we define moment restrictions using the linear pricing rule from not only Regime 3, where both an earnings signal v_F and the FII order flow ω are available, but also from Regime 2, where only the FII trader's order flow ω is available. The extended parametrization in that setup allows us to address questions relating to whether traders behave differently in anticipation of a public announcement.

[12] As a robustness check, we augment the sample of announcements with FII trading with announcements with no FII trades (zeros). Our conclusions are somewhat similar under this variation.

4.2 Centering and a Scale Adjustment

Before estimation, we make two adjustments to UE and one to FIITR. First, in the theoretical model, the information components are zero-mean variables. Therefore, to be consistent, in the empirical counterpart, we mean-center values of UE and FIITR. Second, we also make a scale adjustment to UE. This scale adjustment is made to more easily interpret the coefficient on UE when both UE and FIITR are present, the Regime 3 model.

To understand why and how we implemented the scale adjustment, consider the following. From Proposition 1, the coefficient on UE under Regime 3 is $\beta = 1 + \rho\left(\frac{\sigma_T}{\sigma_F}\right)$. Further, recall that Lemma 1 predicts that the coefficient on UE when there is no FII trading (Regime 1), $\beta = 1$. Thus, the second term in the expression for β measures the indirect effect of FII trading on returns (through UE), when both UE and FIITR are present. But to allow for such an interpretation, we need to transform UE in such a way that its coefficient would equal one, absent FII trading. To do so, we first estimate a linear regression of ERET on UE and control variables for a sample for firm-quarters that has no FII trading during the earnings announcement and no FII ownership (Regime 1). Second, we multiply UE for the Regime 1 sample by the coefficient so obtained and re-estimate the regression; this ensures that $\beta = 1$ for the Regime 1 sample. We then multiply UE for our Regime 3 sample of earnings announcements that have non-zero FII trading by the Regime 1 coefficient estimate. This is our scale adjustment.[13]

4.3 A Remark on Identification; Reparametrization

Our model raises an identification issue that we discuss next. In Proposition 1, inspection of the pricing rule indicates that we have to estimate four deep or primitive parameters – σ_F, σ_T, σ_z, and ρ, and two shallow parameters – β and λ. Note that the three variance-related parameters, σ_F, σ_T, and σ_z, enter the equilibrium solution only in ratio form. This immediately implies that regardless of the estimation criterion we use (e.g., least squares, maximum-likelihood, GMM), we cannot identify all three variance parameters simultaneously. This is because, if a set of values for these three parameters optimizes any given criterion, then so will any scalar multiple of the same values.

[13] Note that, as with OLS regressions, scaling affects the coefficient estimates, but does not affect the t-statistics. We hasten to add that this choice of scale is obviously not a knife-edge choice affecting convergence. A range of scale choices all yield converged estimates.

To address this identification issue, we compute an independent estimate for σ_F from observed values of UE, our proxy for v_F. We equate σ_F, the information advantage of the firm, to the overall sample estimate of the standard deviation of unexpected earnings (UE). Then, in any implementation of the empirical model, we take that independently estimated σ_F as a fixed value and estimate the remaining three primitive parameters. Because β and λ are defined in terms of the deep parameters, they can also be easily computed from the estimates of the deep parameters.

In actual empirical work, we adopt a slight reparametrization of the linear pricing rule in Proposition 1: $p = (\alpha + \beta v_F + \lambda\omega)$, with $\alpha = 0,\ \beta = 1 + \rho \times \sqrt{\sigma_1^2},\ \lambda = \sqrt{\sigma_2^2} \times \sqrt{1 - \rho^2}$. So $\sigma_1 = \left(\frac{\sigma_T}{\sigma_F}\right)$ and $\sigma_2 = \left(\frac{\sigma_T}{\sigma_Z}\right)$. It is important to note that there is strictly no loss of information in this reparametrization, since with an independently estimated σ_F, and estimates of σ_1 and σ_2, $\sigma_T = \sigma_F * \sigma_1$, and $\sigma_Z = \left(\frac{\sigma_T}{\sigma_2}\right)$. Our numerical algorithms search for solutions only within real values, so using the radical guarantees that $\rho \in [-1,1]$.

A subtle but important benefit from adding unexpected earnings in our model should be noted. If we do not have earnings, the same identification problem that we noted above would also apply to σ_T and σ_Z, which would then enter only as a ratio in the definition of λ. In that case, they would be unique only up to a scalar multiple, and only the ratio would be identified. Because we have added unexpected earnings and are able to estimate σ_F independently, we can identify even σ_T and σ_Z, with the simple GMM strategy and our one-period model.

4.4 Motivating and Designing GMM Estimation

There is a large empirical literature related to Kyle's λ, and at least two papers, Foster and Viswanathan (1995) and Cho (2007) even provide estimates of the deep parameters that determine λ. The PIN literature also considers deep parameter estimation. But these papers either ignore model validation altogether, or (as in Foster and Viswanathan (1995)) find that the model is sharply rejected. We also face the challenge of a model that is nonlinear in parameters, and where there is no basis for forming priors about parameters relating to private information. In the context of our model, we overcome this problem by designing a simple GMM strategy that exploits a feature of the pricing rule that has not received attention before in estimation.

Because Regime 3 is a regime with earnings and trading signals, our main model is a model of an event window. Cho (2007) also conditions on earnings announcements. But he does so only to select market data leading up to the earnings announcement on the assumption that in those periods it would be more likely for traders to have private information. He does not use earnings data, and the primitive parameter estimation in his paper and in Foster and Viswanathan (1995) pertains to non-event periods. Our pricing rule involves both earnings and order flows and this, as we have seen, confers some advantages in the identification of the deep parameters. Our moment conditions, while nonlinear in parameters, have a much simpler form than in Foster and Viswanathan (1995) and Cho (2007), who invoke higher-order moments in a dynamic model. This is important for the success of a gradient search algorithm.

The pricing rule is a model of a conditionally expected payoff and a *realized* price, and not of an expected price. This feature of the pricing rule guides our estimation method choice. If we rewrite the model as a pricing error, $u = p - (\alpha + \beta v_F + \lambda\omega)$, under the null hypothesis that the equilibrium model holds, the pricing error $u = 0$. From this, and given the observability of price p, earnings news v_F, and order flow ω, it follows that we can derive moment restrictions for the GMM estimation of the primitive parameters of the model. These restrictions (moment conditions) are functions of the form $h(X, \Theta)$, with $E\big(h(X, \Theta)\big) = 0$, where X denotes data, and Θ the unknown set of k parameters to be estimated. The GMM solution to the estimation problem is obtained by minimizing the quadratic function, $\overline{h(\Theta)}'W\overline{h(\Theta)}$ where $\overline{h(\Theta)}$ is the sample mean of the vector of moment conditions, W is a positive definite and symmetric q x q matrix of weights, where q, $q > k$, is the number of moment conditions.

One class of restrictions that we can employ has the form $E(u'Z) = 0$, where the instrument, Z, is plausibly orthogonal to the pricing error u. Because the equilibrium pricing rule is a model for each realized price, in theory, we could use *any* variable, Z. Of course, there are natural limits to what we can choose as Z. Completely arbitrary variables drawn from say, zoology or oceanography, may help us increase the number of moment restrictions and perhaps even yield lower asymptotic standard errors. But to learn anything useful from a potential rejection of the model or violation of a moment restriction, to begin with, there must be some plausible relationship.

We use only restrictions of the form $E(u'Z) = 0$, and avoid restrictions with higher powers of the error term. Using higher powers of the error would make an already nonlinear model even more nonlinear and add to the complexity of any gradient search algorithm.[14]

While there is a vast literature in finance that applies GMM to the structural estimation of asset pricing models[15], exactly how many and which moment restrictions to select is still at best only an art and not a science. The choice of the number of moment restrictions involves tradeoffs. We need more moment restrictions than parameters to compute an over-identification test statistic to evaluate whether the restrictions are valid. As the number of moment restrictions increases relative to the number of parameters, asymptotically valid GMM standard errors generally shrink, but finite sample performance degrades. Further, using a very large number of restrictions makes model rejection more likely. Our tradeoff is to use five moment restrictions, given that we estimate three parameters. Our five instruments are the constant (1), v_F (UE), ω (FIITR), the pre-announcement week change in the US Dollar-Rupee rate (CH_WK_EXCH), and the pre-announcement week market return (WK_MRET). We elaborate on these instrument choices when we discuss the GMM results in section 6.

A comparison with the moment conditions in Foster and Viswanathan (1995) and Cho (2007) makes clear the relative simplicity of our moment conditions. Together with other elements of our strategy – using a dense grid evaluation of the GMM objective function to obtain good starting points, and a perturbation step to reduce the risk of being trapped in a poor local solution – this lets us find solutions with very low J-statistics for asymptotic validation. We also provide finite sample validation with out of sample model comparisons.

5. Data Sources

We obtain data from four sources: the PROWESS database of the Center for Monitoring Indian Economy Private Limited (Prowess), the website of the Central Depositary Services Limited (CDSL), the Federal Reserve website, and the Thomson Reuters IBES Analyst Forecast database (IBES). Prowess provides the information need to construct the dependent variable (ERET), unexpected

[14] A more general class of restrictions would be $E(u^k Z) = 0, k = 1, 2, 3$ ... In the case of models like the CAPM, the equilibrium relationship involves expected returns; hence, in terms of this notation, what we can take as a primitive is only $E(u) = 0$, not $u = 0$ (see, e.g., Bodurtha and Mark (1991) and Jagannathan and Wang (1996)). Therefore, using powers like u^k in a moment restriction would generally be ruled out.
[15] There is also, of course, a vast literature estimating λ, the shallow parameter in a Kyle model. See e.g., Brennan, Chordia, Subrahmanyam, and Tong (2012). Further, the literature on PIN, the probability of informed trading, has provided estimates not only of PIN but also the primitive parameters determining PIN. See Duarte and Young (2009) and references therein. Our work can be regarded as a counterpart to that just as Glosten and Milgrom (1985) is a counterpart to Kyle (1985). The PIN literature uses the number of buy and sell trades, we use buy and sell volumes.

earnings (UE), and the control variables. The CDSL website is our data source for daily FII buy and sell trades, which we use to construct FIITR.[16] We obtain Dollar-Rupee exchange rates from the Federal Reserve website[17] and Mean Analyst Forecasts of Earnings per share from IBES. We integrate the FII trading data (CDSL) and the firm price and financial statement data (Prowess) by matching on firms' ISINs.

Because we are the first to use the daily FII trading dataset from the CDSL website, we provide a brief overview of it. In this dataset, the basic unit of observation is a trade by an FII for a stock. Data fields include an identifying code for each FII, the ISIN for the stock, the date of the trade, and the exchange on which the trade was executed. In addition, for each trade, the following four variables are available: (a) the number shares bought or sold; (b) an indicator for whether the trade was a buy or a sell; (c) the price at which shares were bought or sold; and (d) the transaction value. Unfortunately, SEBI masks the FII identifying codes and changes the codes every month; consequently, FII-level analysis is difficult. Therefore, for each stock-trading day pair, we aggregate daily data across FIIs. Because we have no reason to expect exchange-related effects, we also aggregate daily trades across exchanges (National Stock Exchange (NSE) and Bombay Stock Exchange (BSE)).[18]

6. Results

6.1. Sample Description

Our sample period consists of fourteen years; it begins in the first quarter of 2003 and ends in the fourth quarter of 2016. The choice of this sample period is dictated by the availability of FII trading data. Our initial sample consists of firms that were listed on the NSE at any point during the sample period. For these firms, only firm-quarters with non-missing earnings announcement dates and quarterly earnings per share were retained. To enter the final sample, firms are required to have non-missing data for unexpected earnings and its four-quarter lagged value, have quarters that end in March, June, September, or December, and announce earnings on dates that are valid (occur after the quarter end) and within 180 calendar days of the quarter-end date. Additionally, firm stock returns should be non-

[16] The data is publicly available on the website of the Securities and Exchange Board of India (SEBI) and some mirror sites. The FII trade data that we employ is from the mirror website maintained by CDSL. Its URL is https://www.cdslindia.com/publications/FII/EquityDataFII.htm.

[17] The URL of our source of exchange rates is https://www.federalreserve.gov/releases/H10/hist/dat00_in.htm.

[18] The CDSL website contains 10,722,359 FII trades during the sample period (2003-2016) for which ISINs matched the ISINs on the Prowess Database. Of these, we retain 10,582,428 trades (98.7%) that were market buys or sells. We exclude non-routine purchases such as purchase of shares in an initial public offering, participation in a rights issue, or shares obtained through conversion of debentures (140,111 trades). We also exclude 13,808 trades (0.13%) that were executed on exchanges other than the NSE and the BSE. The latter are the largest and most liquid exchanges in India. Our measures of daily net FII buying are based on the remaining 10,568,620 trades.

missing for at least 45 days during the 90 trading-day period centered on the earnings announcement date and for the two earnings announcement days. We also exclude firm-quarters that are missing data for any of the control variables and firms with only one observation over the sample period.[19] In our final sample, 17,877 firm-quarters (30%) had FII trading during the earnings announcement window (Regime 3); 24,101 firm-quarters (40%) had some FII ownership before or after the earnings announcement, but no FII trading during the earnings announcement (no-trade sample); and 18,078 firm-quarters (30%) relate to firms with no trading during the earnings announcement period and no FII ownership both before and after the earnings announcement (Regime 1).[20]

The bulk of our analysis is based on the Regime 3 sub-sample (17,877 firm-quarters). We use the Regime 1 sub-sample to mainly obtain the scaling adjustment that is applied to the Regime 3 sample. Additionally, we examine the robustness of our results to the inclusion of firm-quarters with no FII trading during earnings announcements. To do so, we combine the Regime 3 sub-sample with the zero-trade firm-quarters from the other two sub-samples that relate to the firms from the Regime 3 sub-sample (39,346 firm-quarters).

Table 1 presents univariate statistics for the pricing model variables for the Regime 3 sub-sample. All variables are winsorized at 1% and 99% levels. The mean earnings announcement return (ERET) is -0.09%, but the median is -0.32%, suggesting the influence of some large positive values on the mean. Mean unexpected earnings scaled by share price (UE) is negative at –0.18%; however, the median is slightly positive at 0.10%. The mean and median net FII buying (FIITR) at the earnings announcement is almost zero. The average zero net buying masks the fact that FIIs are buying and selling on that day, and their buys and sells offset each other.[21]

Turning to the control variables, Table 1 reports that the mean market return (MRET) is positive at 0.04%, the log of mean market capitalization (LMCAP) is 10.24, and the mean book-to-market ratio (BM) is 0.59. Mean return in the fiscal quarter before the earnings announcement (MOM3) is positive, on average (9.5%). The sample firms are profitable on average and are growing – mean operating profitability (OPROF) is 18.99%, and mean asset growth (AGRO) is 22.33%. Mean monthly return

[19] The 45-day non-missing return requirement is imposed to ensure that our sample firms are fairly liquid. Having at least two observations per firm over the sample period eliminates singleton firms. We estimate panel regressions and singletons will cause standard errors to be underestimated.

[20] Table IA.1 in the Internet Appendix presents the filters applied to arrive at the final sample of 60,056 firm-quarters. Table IA.2 in the Internet Appendix reports the proportions of three sub-samples over the sample period. The frequencies indicate that, except for 2003 and 2004, the relative proportions of the three firm-quarter types do not display a significant temporal shift during the sample period.

[21] Figure IA.1 in the Internet Appendix shows how median FII buying, selling, and net buying behave around earnings announcements. On days 0 and 1, both FII buying and selling spike to close to 0.04%. Median levels on these days are higher than any other day in the sixty-day window around earnings announcements.

volatility (STDRET) is 2.64%, and the mean monthly volume as a percentage of shares outstanding (VOL) is 27.98%. The average logarithm of age (L_AGE) is 2.59, average dividend yield (DIVY) is 1.56%, and the average logarithm of price (LPRC) is 5.40. Table 1 also reports distributional information for two variables that we employ as instruments in the GMM (and 2SLS) estimation – the change in the US Dollar-Rupee rate over the week ending on day -1 relative to the earnings announcement period (CH_WK_EXCH) and the market return over the week ending on day -1 (WK_MRET).

Table 2 reports simple correlations between all our variables. ERET is positively and significantly related to both UE (0.06) and FIITR (0.09), suggesting that both are priced. The simple correlation between UE and FIITR is -0.0015 and not significantly different from zero. Recall from the discussion in section 3.3 that our model implies that the two public signals will be orthogonal to each other. The data appears to be consistent with this implication of the model. Note, however, that despite the orthogonality of UE and FIITR, the correlation between the *information* of the firm and FIIs could be non-zero. While the model estimates of the deep parameter ρ will reveal this commonality, the simple correlation reported in Table 2 does not.

6.2. Regression models

Our analysis of the pricing of UE and FIITR begins with the benchmark case when UE is the only signal available to the market (Regime 1). We estimate a panel regression of ERET on mean-centered UE, control variables, firm effects, and year effects for this sub-sample (n=18,078). The untabulated coefficient estimate on UE in this regression is 0.025542. We re-estimate the Regime 1 regression with mean-centered UE multiplied by this scale factor and obtain a coefficient of exactly one for centered UE. Before estimating regressions of ERET on both UE and FIITR for the Regime 3 sample, we multiply the centered UE for all observations in that sample by 0.025542.

In Table 3, we report four OLS regressions for the Regime 3 sample. The dependent variable is ERET, and the main independent variables are UE and FIITR, whose coefficients are β and λ, respectively. We include firm and year fixed effects and adjust standard errors for clustering within each firm. To reduce the impact of outliers, we winsorize ERET, UE, FIITR, and MRET at the 1% and 99% levels, by year. All control variables are first transformed into decile ranks ranging from 1 to 10 and then transformed again such that their values lie on the [0, 1] interval; we do so by subtracting one from the ranks and dividing the remainder by nine. Note that these regressions do not use any restrictions from the underlying theory, and the coefficients β and λ here are estimated directly from the data, without invoking any of our equilibrium formulae.

Column (1) reports the baseline regression that includes only control variables. The adjusted R^2 for this regression is 14.88%. In column (2), we augment the model with UE. The coefficient on UE is 9.037, with a p-value of 0.00, and the adjusted R^2 increases to 16.77%. In the regression in column (3), we include FIITR but exclude UE. FIITR is positively and significantly related to ERET and has a coefficient of 5.219 (p-value = 0.00). The adjusted R^2 in this regression is 18.13%, which compares favorably to that of the regression in column (2) and suggests that FIITR is a more influential determinant of returns than is UE. Column (4) reports the Regime 3 regression results when both UE and FIITR are included. Compared to the results in columns (2) and (3), the coefficients β and λ are essentially unchanged and statistically significant, and the adjusted R^2 climbs to 20.08%. The stability of β and λ across regressions confirms that the variables UE and FIITR are independent of each other. Neither β nor λ increases or decreases because of the presence of the other signal.

A critical assumption underlying the model estimates in columns (1) to (4) of Table 3 is that FII trading is exogenous to earnings announcement returns. This is in contradiction to our theory, where we endogenize FII trading. Additionally, it is very plausible that FII trading responds to price movements during the earnings announcement period. To account for the endogeneity of FIITR, we employ two-stage least squares (2SLS) and re-estimate the ERET regressions. Our instruments for FIITR are the change in the US Dollar-Rupee rate over the week ending on day -1 relative to the earnings announcement period (CH_WK_EXCH) and the market return over the week ending on day -1 (WK_MRET). We chose these instruments as we expect them to be unrelated to firm-specific news during the earnings announcement period, but likely correlated with FII trading. The two instruments are also winsorized at the 1% and 99% level, by year.

Table 3, Column (5) reports the 2SLS results. At the bottom of Table 3, we report diagnostics related to the validity of the two instruments. The Durbin-Wu-Hausman test statistic for endogeneity is 4.06 (p-value = 0.04), implying that the null hypothesis of FIITR exogeneity is rejected. In untabulated first-stage regressions of FIITR on the two instruments and all the other exogenous variables (UE, control variables, firm effects, and year effects), both instruments are significantly related to FIITR. The t-statistic on the CH_WK_EXCH is 1.95, and that on the WK_MRET is 4.66. Importantly, these instruments are not weak; the Kleibergen-Paap rank Wald F statistic is 10.99, which exceeds the Stock-Yogo Weak ID 10% critical value (8.68). Additionally, the Hansen-Sargan J Statistic for over-identification is 1.42 (p-value = 0.23). Thus, the two instruments are likely uncorrelated with the error term in the earnings announcement return regression. Overall, our diagnostics suggest that the two instruments are valid.[22]

The 2SLS results when we account for the endogeneity of FIITR indicate that coefficients on UE and FIITR continue to be significantly related to ERET. The coefficient on UE, β, is 9.437 (p-value = 0.00), which is slightly larger than that obtained under OLS in column (4). The coefficient on FIITR, λ, is 17.602 (p-value = 0.01), which is about 3.3 times its value under OLS. Overall, our conclusions about the effect of UE and FIITR on prices are largely unchanged when we account for the endogeneity of FIITR. But the magnitude of the impact of FIITR is larger.

6.3 Estimates of Primitive Parameters

Thus far, to obtain estimates of β and λ, we have not accounted for the model structure predicted by Proposition 1. The model is defined by the pricing error $u = p - (\alpha + \beta v_F + \lambda\omega)$, with $\alpha = 0, \beta = 1 + \rho * \sqrt{\sigma_1^2}, \lambda = \sqrt{\sigma_2^2} * \sqrt{1-\rho^2}$, where $\sigma_1 = \left(\frac{\sigma_T}{\sigma_F}\right)$ and $\sigma_2 = \left(\frac{\sigma_T}{\sigma_Z}\right)$. Here, ρ is the correlation between the information of the firms and the FIIs, σ_1 is the ratio of the information advantage of the FIIs (σ_T) to that of the firm (σ_F), and σ_2 is the ratio of the variance of informed trading (σ_T) to the variance of noise trading (σ_Z). Our goal is to estimate the three primitive parameters of the pricing model: ρ, σ_1, and σ_2.

Our GMM approach to estimate the pricing model is to specify moment conditions of the form $E(u'Z) = 0$, where the instrument, Z, is plausibly orthogonal to the pricing error u. We chose five instruments that are in turn the constant (1), v_F (UE), ω (FIITR), the pre-announcement week change in the US Dollar-Rupee rate (CH_WK_EXCH), and the pre-announcement week market return (WK_MRET). If our model does not have any correlated omitted variables, UE and FIITR would be uncorrelated with the pricing error. The Hausman test results from our 2SLS estimation suggest that both CH_WK_EXCH and WK_MRET are likely uncorrelated with the pricing error. Because the number of moment conditions exceeds the number of parameters, we report a J-test to verify if the moment conditions are asymptotically valid.

[22] The instrument diagnostics were obtained by using xtivreg2 command in Stata. We report the Kleibergen-Paap Wald rank F statistic for weak instruments because, unlike the Cragg Donald F statistic, the former is robust to violations of the i.i.d error assumption. Again, because the i.i.d. assumption may not be reasonable, the estimates reported in Table 3, column (5) are based on the continuously updated GMM estimator, or CUE estimator.

To obtain parameter estimates, we use the Iterated GMM (IGMM) procedure suggested by Hansen (1982). Because the objective function is nonlinear in parameters, we need to compute GMM estimates by numerically minimizing the objective function; this would require a careful choice of starting values. To find good starting values we evaluate the GMM objective function at a dense grid, exploiting the mathematical structure of the problem to make the computer intensity feasible. Further, because our objective function is non-convex, we adapt a perturbation proposed by Wood (2001) to the IGMM procedure. The Wood perturbation reduces the chance of being trapped in a flat region of the objective function and makes it more likely that we move to a better local minimum. Appendix C contains a summary of the steps we employ to define starting values and obtain parameter estimates.[23] The standard errors of our parameter estimates are bootstrap standard errors that account for heteroscedasticity. Again, Appendix C describes the bootstrap methodology to compute standard errors.[24]

In Table 4, Column (1), we report the GMM model estimates of the three primitive parameters and bootstrap standard errors. The p-values show that all parameters are significant at conventional levels of significance. The striking result from Table 4 is in the magnitudes of σ_1 and σ_2. The σ_1 parameter estimate of 23.75 suggests that the FIIs' information advantage, σ_T, dwarfs the firm's information advantage, σ_F. This does not mean that the *total* information in the firm's report is small. That could still be large, but communication throughout the year, or close following by analysts of these relatively larger firms, could ensure that most of that information is already known to others. The σ_2 parameter estimate of 4.72 suggests that what FII traders know is not dwarfed by background noise. The parameter ρ is 0.249, indicating that firms' and FIIs' information share a common component. We also use the formula in Proposition 1 to calculate the values of β and λ. Table 4 reports that they are respectively, 6.908 and 4.575. Recall from Lemma 1 that the benchmark estimate of β when there is no FII trading is 1. The value of 6.908 for β when FIIs do trade suggests that the market learns about what traders know from the firm's earnings report. Also, the GMM estimates of β and λ are lower than those obtained under 2SLS (9.463 and 18.073, respectively). Lastly, the Hansen J-Statistic equals 0.002 (p-value = 0.99). Hence, the results asymptotically support the null hypothesis that the model is valid.

For completeness, Table 4 reports estimates of σ_F, σ_T, and σ_Z as well. From the pricing rule, $\sigma_1 = \left(\frac{\sigma_T}{\sigma_F}\right)$ and $\sigma_2 = \left(\frac{\sigma_T}{\sigma_Z}\right)$. Hence, by fixing σ_F, our estimate of the information advantage of the firm,

[23] Additional details are in Internet Appendix IA.B.

[24] As an alternative, we also compute wild *cluster* bootstrap standard errors that account for within firm serial correlation in errors. For the results reported in Table 4 we find that the wild bootstrap standard errors are approximately 98% of the size of the wild cluster bootstrap standard errors. Since are inferences are very similar for the two procedures, for the remaining tables in the paper, we report results based on the simpler wild bootstrap.

we can back out σ_T and σ_Z. We compute σ_F as the standard deviation of unexpected earnings (UE). To compute its standard error, we resample from the UE vector 1,000 times to construct 1,000 bootstrapped UE vectors. We then compute the standard deviation of each resampled UE vector. The standard deviation of these standard deviations is the standard error of σ_F. As with the other deep parameters, the results indicate that σ_F, σ_T, and σ_Z are statistically significant at conventional levels.

To the best of our knowledge, papers that analyze institutional trading have tended to focus only on periods with such trading, so that the estimates are really conditional on the existence of trading. Easley, Kiefer, O'Hara, and Paperman (1996) suggest that non-trading can also be an important signal. This means that even if we do not have FII trading data in our original source file, rather than coding it as "no data" we should code it as a "no-trade" or "zero-trade." To evaluate the impact of no-trades on our results, we augment our Regime 3 sample with earnings announcements during which there was no FII trading. No-trades relate only to firms that belong to the regime 3 sample.[25] The augmented sample size is 39,346, which is more than twice the Regime 3 sample. Column (2) of Table 4 reports the parameter estimates when we include the "no-trades." As we should expect, with the large number of no-trades, and the consequent dampening of variation in FII trades, the estimates for σ_1(7.75), and so of σ_T (0.018) are lower. But the other qualitative features of our estimates hold: ρ is positive (0.282), $\sigma_2 > 1$ (5.749), and $\beta > 1$ (3.183). The effect of FII trading on price is larger with the no-trade augmentation; λ equals 5.516. The J-statistic continues be small (p-value = 0.99), implying model validity.

6.3.1 Model Validation

The dominant tradition in the empirical literature relating to Kyle's λ and PIN has been to assume the underlying model is valid. So Cho (2007) focuses on testing whether models nested within the Foster and Viswanathan (1995) framework can be rejected, without testing the Foster-Viswanathan model itself. Foster and Viswanathan (1995) do test their model and show that it is sharply rejected. The Hansen J-statistic in Table 4 is small and consistent with the underlying model being valid. The GMM literature has, however, noted that the p-value of the J-Statistic is only asymptotically valid and can be a poor guide to the true error rate in a finite sample. So, to further address model validity, we implement a ten-fold cross-validation exercise. The steps in this exercise are as follows:

[25] In Table 1 we report that 'no trades" consist of 42,179 (24,101 + 18,078) observations. About 50.9% of these 42,179 observations (21,469) relate to firms in the regime 3 sample. By drawing "regime 3" firms from the no trade samples, we attempt to reduce the likelihood that a characteristic related to the "trade versus no-trade" decision influences our results.

1. Randomly divide the Regime 3 sample (17,877 obs.) into ten groups.
2. Treat one group as the holdout or test dataset (1,787 obs.).
3. Treat the remaining groups as the training dataset (16,090 obs.).
4. Estimate the model on the training set and use the parameters obtained to predict the dependent variable (ERET) for the test dataset.
5. Compute forecast errors as the difference between realized ERET and predicted ERET for the test dataset and then compute the Root Mean Squared Error (RMSE) and the Mean Absolute Error (MAE).
6. Repeat steps 2 to 5 for each of the remaining nine groups.

We implement the cross-validation procedure using both OLS and GMM. Table 5 reports RMSE and the MAE for the ten groups for forecasts based on OLS and the GMM in columns (1)-(2) and columns (5)-(6), respectively. The results indicate that the forecast accuracy of the two models is very similar. So, while the application of the underlying equilibrium model yields information about primitive parameters, the out-of-sample performance of the GMM model is not significantly better than the purely statistical OLS model. Table 5 also reports p-values from a t-test (Wilcoxon Signed-Rank test) of whether the mean (median) forecast error equals zero. The t-test results (columns (3) and (7)) indicate that for every group, the null of zero-mean forecast errors cannot be rejected at conventional significance levels. The p-values for the t-test range from 0.37 to 0.93. For the non-parametric Signed-Rank test, the p-values range from 0.01 to 0.33. Note that to achieve a family error rate of 1%, by the Bonferroni criterion, the individual test Type I error rate should be no more than 0.1%. So even with the Signed-Rank test, the null of zero-median, and therefore model validity, cannot be rejected.

6.3.2 Firm-by-firm analyses

To assess if there is heterogeneity in primitive parameter estimates, we estimate time-series regressions of the pricing model for individual firms. To enter this analysis, we require that a firm have at least twenty quarters of data during the sample period. This screen results in a sub-sample of 365 firms compared to the Regime 3 sample of 1,132 firms. These 365 firms span 12,118 firm-quarters (the firm-by-firm sub-sample).

In Table 6, we summarize the distribution of estimates for the 365 firms. The mean ρ is 0.321, and the median is 0.795, suggesting that the distribution of ρ is left-skewed. Further, in untabulated findings, $\rho < 0$ for 114 firms, implying that FIIs and firms have the opposite information about future payoffs for a significant fraction of the sample. The mean values of σ_1 and σ_2 are positive and well above one (29.659 and 22.381, respectively). These values confirm that FIIs' information advantage is, on average, significantly larger than that of the firm, and FII trading during earnings announcements is more news than noise.

The shallow parameters' estimates - β and λ, are both positive and significant on average (13.666 and 8.343, respectively). Interestingly, $\beta < 0$, for 102 firms or 27.9%, of the 365 firms. This occurs when ρ < 0 and $\sigma_1 \gg 0$. Thus, the combination of a significant FII information advantage and a negative correlation between firms' and FIIs' information can cause good news to be interpreted as bad news. Note also that when ρ = –0.99 or +0.99, λ becomes very small. This occurs about 16.2% of the time (39 firms at the upper bound and 20 at the lower bound). When ρ takes on extreme values, the market anticipates all of the trader's private information, and no information advantage remains. Lastly, in terms of model evaluation, for the 365 firms, Hansen's J statistic ranges from 0.001 to 0.692 with associated p-values that range from 0.774 to 0.994. Thus, the model fits the data well for all the firms.

The last column of Table 6 (column (8)) contains aggregate estimates for the sub-sample used in this firm-level analysis. This sub-sample has qualitatively similar characteristics to the main sample in Table 4 (n = 17,877) in terms of β, λ, and σ_F. However, the $\rho = 0.773$ for the sub-sample is much higher than that of the main sample ($\rho = 0.321$). Also, while σ_1 for the sub-sample is lower than that of the main sample, σ_2 is larger.

6.3.3 Firm and Trader Characteristics and Primitive Parameters

In Table 7, we dig a little deeper into our primitive parameter estimates. Our first partition of the data is based on firm size. Prior research on the pricing of earnings has interpreted firm size as a measure of the information environment of the firm (Collins, Kothari, Rayburn (1987)). Earnings announcements are generally viewed as more informative for smaller firms (larger β) because there is less pre-announcement information for these firms (Ball and Shivakumar (2008)). We measure size as the log of market capitalization at the beginning of the quarter (LMCAP). To form size groups, we divide the sample into ten deciles based on LMCAP. Firm-quarters in smallest three deciles are small firms, those in the next four deciles are medium firms, and those in the largest three deciles are large firms.

Panel A of Table 7 reports parameter estimates across size groups. The parameter β, while always greater than 1, is not monotone in firm size. It is the highest for medium-sized firms. This is explained by the elements of $\beta = 1 + \rho\left(\frac{\sigma_T}{\sigma_F}\right)$. Panel A reveals that both the numerator, $\rho \times \sigma_T$, and the denominator, σ_F, decline in firm size, but at different rates, resulting in a nonlinear relation between announcement returns and UE. The negative relation between σ_F and firm size is consistent with prior research that larger firms have a greater demand for information gathering and thus less of an information advantage. However, prior research has not considered the effect of firm size on $\rho \times \sigma_T$, the information advantage of strategic traders. Thus, we provide a more nuanced view of how firm size affects the pricing of earnings.

The coefficient on FIITR, λ, measures the signal-to-noise ratio associated with FII trades. It has in its numerator, $\sigma_T * \sqrt{1-\rho^2}$, i.e., the trader's gross information advantage σ_T adjusted by ρ, a measure of how much of the trader's private information can be guessed from just the firm's announcement v_F. The denominator σ_Z measures market noise. While λ increases with firm size, this is a result of the complex interaction of the primitive parameters - σ_T rises slightly with size, even as σ_Z declines. But ρ declines more sharply with size, suggesting that for very large firms we can glean less from the firm's report about the traders' private information. This is not inconsistent with the traditional belief that more is known about large firms. There may be more *total* information about large firms (e.g., simply from more pre-announcement information-gathering and a larger analyst following) yet traders may have more of an information advantage with respect to these corporations because they pay more attention and allocate more resources to following them, as they have larger stakes in them. This finding does suggest that we become more sensitive to the distinction between total information available and a trader's information advantage and be cautious before using firm size as an information asymmetry proxy.

Our second partition is based on the sign and size of earnings per share. Previous work (e.g., Hayn (1995)) has noted that β for loss firms is less than that for profitable firms. Hayn (1995) argues that given a liquidation option, losses are less likely to persist and are less informative about future prospects causing them to be valued less. The magnitude of earnings has also been shown to matter for valuation. Beginning with Burgstahler and Dichev (1997), several studies document a sharp discontinuity around zero for various profit measures, with a disproportionate number of firms reporting profits just to the right of zero. These "small profit" firms are regarded as firms that have just managed to avoid reporting a loss by managing earnings upward. Given their lower earnings quality, "small-profit firms" are expected to be valued less.[26]

Panel B of Table 7 reports the deep and shallow parameter estimates for the three groups – small-profit firms, loss firms, and larger-profit firms. In column (1), the parameter β is negative, although statistically insignificant, for small-profit firms. This can be explained by $\sigma_T > \sigma_F$ and $\rho < 0$. The relative information advantage of the FIIs over that of the firm is the highest for this group. Further, the firms and FIIs disagree considerably on the interpretation of the information that is common to them. Consequently, the small amount of 'good news' is perceived as bad news by investors. Interestingly, the signal-to-noise ratio, λ, is highest for firms with small profits.

As columns (2) and (3) indicate, consistent with prior research, the β for loss firms is less than that for larger-profit firms. Our deep parameter estimates are consistent with a different albeit complementary explanation for the valuation differences between the two groups of firms. Bad news is more closely held, so to learn more, let alone learn more so that a trader will have an advantage over others, is more difficult, leading to a lower σ_T for loss firms. While loss firms' ρ is higher than that of larger-profit firms, as σ_T is lower, the product $\rho \times \sigma_T$ (the numerator of β) is only slightly larger for loss firms. Ceteris paribus, this would cause a higher β for loss firms. But because loss firms also have a greater information advantage (σ_F) than larger-profit firms, the weight on v_F in the expression for $E(\tilde{v}_T \,|v_F)$, and therefore β, is less.

The signal-to-noise ratio relating to FII trades, λ, is larger for large-profit firms compared to that of loss firms. While the amount of noise trading for profitable firms is more than that of loss firms, the FII information advantage for the former group is significantly larger than that of the latter; consequently, FII trades are more informative for profitable firms.

In Panel C of Table 7, we examine how FIIs' attention during earnings announcements affects their absolute and relative information advantage. We assume, as is the tradition in the literature on attention (e.g., Hirshleifer, Lim, and Teoh (2009)), that a trader's attention is divided among competing simultaneous earnings announcements. In panel C, attention is measured using the average number of market-wide earnings announcements over days [0, 1], and firms are divided into three groups – bottom

[26] Internet Appendix Figure IA.2 presents a partial histogram of earnings per share around 0, ranging from ₹-2.0 per share to ₹2.0 per share with a bin width of 0.1 for the main sample (n = 17,877). Consistent with the U.S. evidence, the figure indicates a sharp discontinuity around zero, with a disproportionate number of firms to the immediate right of 0, compared to the number of firms to the left of zero. Loss avoiders, or small-profit firms, are defined as firms with quarterly earnings per share (EPS) that is > ₹0.00 and ≤ ₹0.01.

30% (high attention), middle 40% (medium attention), and top 30% (low attention). Attention is assumed to decrease as the average number of market-wide earnings announcements over days [0, 1] increases. The results indicate that, as one would expect, both the absolute (σ_T) and relative information advantage (σ_1) of the FIIs increases as attention decreases. Because we have no priors on the other parameters, we just note that many features of the main sample ($\rho > 0$, $\sigma_1 > 1$, $\sigma_2 > 1$) hold across sub-samples.[27]

6.3.4 Using analyst forecast data

A substantial body of research in accounting and finance uses mean analyst forecasts of earnings per share to measure pre-announcement expectations. In our main results, discussed thus far, we use lag-4 quarterly earnings per share to measure this expectation. Because analysts can incorporate information releases over four quarters, their forecasts are potentially more accurate. However, several researchers have argued and provided evidence that analyst forecasts are biased because of behavioral or incentive-related reasons. Thus, which of the two measures is a better proxy for earnings expectation is an open question.

Unfortunately, analyst forecast data is very limited in India. This lack of data is the reason why we used lag-4 quarterly earnings per share as the expectation for the analyses thus far. Of the 17,877 earnings announcements, we were able to obtain only 910 announcements with mean analyst forecasts on IBES (hereafter analyst sub-sample). For an observation to enter the analyst sub-sample, we require that its mean forecast be based on least three individual forecasts, and each of these forecasts should have been issued in the month before the earnings announcement to reduce staleness.[28]

In Table 8, we report model estimates for the analyst sub-sample using the two alternate proxies for earnings expectations. Column (1) contains the estimates based on lag-4 quarterly earnings per share as the earnings expectation to compute unexpected earnings (UE). The qualitative features of the estimates are similar to those obtained for the full sample in Table 4 - $\rho > 0, \sigma_1 > 1, \sigma_2 > 1$. In terms of magnitudes, ρ for the analyst sub-sample is about twice that of the main sample, σ_T is thrice that of

[27] As an alternative to the number of competing earnings announcements, we also construct a measure of attention based on the average number of other firms that FIIs trade during the earnings announcement window. More trading in other firms is likely to reduce attention. The estimates based on this alternate measure show no clear effects on σ_1 or σ_T.

[28] We also considered using analyst forecast data from Bloomberg. But the difficulty in using that data is that the individual estimates that go into calculation of the mean are not separately observable. Therefore, there is no way to tell how many forecasts enter the calculation of the mean or when they were issued.

the main sample, and σ_Z is about one-fifth that of the main-sample. Thus, the magnitudes in the analyst sub-sample are different from the main sample.

In column (2), when we use mean analysts' forecast as expectation, ρ almost doubles relative to that obtained from using four-quarter lagged earnings per share; it equals 0.981. Thus, FIIs and firms appear to share a significant amount of information related to future payoffs when we use analyst forecasts to define expected earnings. Because analyst forecasts are much more accurate than historical earnings-based forecasts, the standard deviation of unexpected earnings based on the former is much smaller, only one-tenth. Consequently, even though σ_T is equal for the two models, σ_1 and hence β are significantly larger when analysts' forecasts are used as earnings expectations. Lastly, σ_Z for the model using analyst forecast as expectations is about one-fourth that using historical earnings; this results in σ_2 being about four times larger for the analyst-forecast-based model.

Because the analyst sub-sample could reflect special characteristics, we do not want to over-emphasize its implications. Our findings for this sub-sample confirm that (a) firm and FII information sets are positively correlated; (b) markets perceive FIIs as being more informed than firms, and (c) informed trading during earnings announcements is not dwarfed by noise.

6.3.5 Earnings and trading signals, substitutes or complements?

Thus far, the analysis has focused on the pricing of UE and FIITR when both signals are present, the Regime 3 model. By comparing estimates when both signals are available with estimates when only one signal, either UE or FIITR, is available, we can assess whether the two signals are substitutes or complements, or if they are independent. When there are two signals X and Y, if the weight on X increases in the presence of Y, then Y is an information complement to X. If the weight on X decreases in the presence of Y, then Y is an information substitute for X. Else X and Y are independent.

To assess the interdependence of UE and FIITR, we jointly estimate the model where both signals are present (Regime 3) and the model where only FII trading is present (Regime 2).

The Regime 3 model expressed as an equation for the pricing error is:

$$u_1 = ERET - (\beta^{FT} * UE + \lambda^{FT} * FIITR) \quad (2)$$

with $\beta^{FT} = 1 + \rho * \sqrt{\sigma_1^2}$, $\lambda^{FT} = \sqrt{\sigma_2^2} * \sqrt{1-\rho^2}$, where $\sigma_1 = (\sigma_{T1}/\sigma_F)$ and $\sigma_2 = (\sigma_{T1}/\sigma_{Z1})$.

Note that the coefficients of UE and FIITR have been modified with a superscript "FT" to highlight the presence of the two information components (firm and trader). Additionally, the parameters σ_T and σ_Z have been modified as σ_{T1} and σ_{Z1}, respectively, to distinguish the Regime 3 parameters from the Regime 2 parameters (described below).

The Regime 2 model, defined in Lemma 2, describes the pricing of FIITR in non-announcement periods. From Lemma 2, the empirical specification of the pricing error when only FII trading is present is given by:

$$u_2 = NON\_ERET - (\lambda^T * NON\_FIITR)$$

(3)

where NON_ERET is the non-announcement period return, NON_FIITR is the non-announcement period FII trading, and $\lambda^T = \sigma_3 = (\sigma_{T2}/\sigma_{Z2})$

In Equation (3), the coefficient on NON_FIITR has the superscript "T" to indicate that only one information component related to the strategic trader is present. The measurement of NON_ERET and NON_FIITR in Eq. (3) requires a choice of a non-announcement date. The period immediately before the earnings announcement is somewhat unique because of insider trading restrictions and the issuance of earnings guidance. Hence, we define days [-31, -30], which is about six weeks before the announcement, as the non-announcement period and thus abstract from pre-announcement effects.

The joint estimation of Eq. (2) and (3) requires some thought about which parameters can be identified. One alternative would involve the estimation of four parameters: σ_1, σ_2, σ_3, and ρ. Unfortunately, with this specification, we cannot solve for the Regime 2 deep parameters, σ_{T2} and σ_{Z2}. The most general specification would involve estimating five parameters: σ_{T1}, σ_{T2}, σ_{Z1}, σ_{Z2}, and ρ. But this creates an identification problem because the solutions for σ_{T2} and σ_{Z2} are valid only up to a scalar multiple. Therefore, we pursue two specifications where we impose a cross-regime restriction to identify the deep parameters in Regime 2.[29] In the first specification, we require $\sigma_T = \sigma_{T1} = \sigma_{T2}$ and estimate the four parameters σ_T, σ_{Z1}, σ_{Z2}, and ρ (the common σ_T specification) and in the second specification, we require $\sigma_Z = \sigma_{Z1} = \sigma_{Z2}$ and estimate the four parameters σ_{T1}, σ_{T2}, σ_Z, and ρ (the common σ_Z specification).

[29] Note that even for the two specifications where the primitive parameters for Regime 2 cannot be identified, the shallow parameter λ^T is identified. Results from those specifications are available from the authors. They yield similar qualitative results in the assessment of earnings or trading signals as substitutes or complements. In particular, $\lambda^{FT} > \lambda^T$, even under those specifications.

The GMM joint estimation of Eq. (2) and Eq. (3) requires specifying moment conditions for the two equations. In addition to the five moment conditions associated with regime 3, we employ four additional moment conditions involving the product of the pricing error of the regime 2 model, u_2 and the following instruments: 1 (constant), NON_FIITR, the change in the US Dollar-Rupee rate over the week ending on day -29 relative to the earnings announcement period, and the market return over the week ending on day -29. Thus, we have nine moment conditions in all, five associated with Regime 3, and four with Regime 2. Internet Appendix IA.A summarizes the specification of the pricing errors, u_1 and u_2, instruments employed to define moment conditions, estimable parameters, and parameter restrictions.

Table 9 contains the parameter estimates from the joint estimation of the Regime 3 and Regime 2 models. Column 1 reproduces the deep and shallow parameter estimates from the baseline specification based only on Regime 3 (see Table 4). The next two columns provide estimates for the common σ_T and common σ_Z model, respectively. Several qualitative features hold across the three specifications. The J statistic is always very small, so model validity is not rejected. Also, $\sigma_1 \gg 1$, $\sigma_2 \gg 1$, $\beta^{FT} > 0$, and $\rho > 0$. The values of λ^{FT} and λ^{T} are stable across specifications.

To understand whether the two signals are substitutes or complements, first note that if the deep parameters are constant across regimes, then $\lambda^{FT} \leq \lambda^{T}$, with equality only when $\rho = 0$. So except for when $\rho = 0$ (when the earnings signal v_F and the FII private information v_T are independent), it would always be the case that for price-setting market makers, the earnings signal v_F is an information substitute for the FII trading signal ω. However, the converse is not necessarily true. Whether the FII trading signal is a complement to, or independent of, or a substitute for v_F, depends crucially on whether $\rho >, =,$ or < 0, as this affects whether $\beta^{FT} >$, $=$, or $< \beta^{F} = 1$. As a practical matter, the primitive parameters are not constant across regimes.

Since $\beta^{FT} = 6.9 > \beta^{F} = 1$ and $\lambda^{FT} = 4.6 > \lambda^{T} = 4.4$ in both the two-regime specifications in Table 9, the presence of each signal increases the weight on the other. So FII trading and earnings are mutual complements. The reason for complementarity is not a confirmation effect, as suggested by Gonedes (1978) and Allen and Ramanan (1990). Rather, FII trading is a complement to earnings because the response to earnings also includes a response to anticipated information of traders. And earnings is a complement to FII trading because the reduction in the trader's information advantage when earnings are also available, is less than proportionate to the reduction in market noise. This result also complements the results in Hirshleifer, Lim, and Teoh (2009) that traders' have finite resources of attention that they allocate among firms that make simultaneous announcements.[30]

In this paper, our two regimes are static and independent, rather than being two periods of one dynamic model. A dynamic model in the spirit of Foster and Viswanathan (1995) may be a useful exercise. But these results suggest that if we had a two-period model, we might see some evidence consistent with anticipation of a public announcement.

7. Conclusions

We build a Kyle-type pricing model with earnings and trading signals. The primitive parameters are the relative information advantages of traders and firms, the correlation between the information of firms and traders, and the variance of noise trades. The central innovation in our GMM strategy is the use of moment conditions derived from the equilibrium pricing rule.

The results for our sample drawn from Indian data suggest that traders know more about firm payoffs than firms themselves. The reaction to earnings is as large as it is because market participants are also using earnings to learn about the private information of traders. On average, firms' and traders' information are positively correlated. We also find that, after accounting for endogeneity of trading, the information contained in trades exceeds the noise.

We find that for many firms in our sample (about 28%) a combination of (i) firms and traders disagreeing about a piece of information related to future payoffs and (ii) the trader's informational advantage being sufficiently larger than that of the firm causes the market's weight on unexpected earnings to be negative. So good news about firm earnings can be viewed as bad news by markets as noted in a different setting by Lundholm (1988) and Manzano (1999). The traditional result that unexpected earnings are weighed positively may reflect the omission of a key market signal, institutional trades. This conclusion is possible only because we explicitly model the underlying equilibrium in a correlated environment and confront that model with data.

The simplicity afforded by the component payoff structure can be useful in other applications. Component structure meets the test of Occam's Razor. It provides the simplest explanation of why

[30] This motivated the alternative definition of trader attention in footnote 28 of this paper, which focuses on the number of different stocks a trader trades in on a given day, rather than looking at the number of simultaneous announcers, as prior work has done.

earnings do not sufficiently account for the price reaction even within an earnings announcement window. There are significant other components of payoff, some not directly observable even to firms, that we are yet to identify.

References

Acharya, V., R. Anshuman, and K. Kumar. 2014. Foreign fund flows and stock returns: Evidence from India. Working Paper, New York University.

Admati, A., 1985. A noisy rational expectations equilibrium for multi-asset securities markets. *Econometrica* 53: 629-658.

Admati, A. and P. Pfleiderer. 1988. A theory of intraday patterns: Volume and price variability. *Review of Financial Studies* 1:3-40.

Allen, S. and R. Ramanan. 1990. Earnings surprises and prior insider trading: Tests of joint informativeness. *Contemporary Accounting Research* 6:518-543.

Ball, R. and L. Shivakumar. 2008. How much new information is there in earnings? *Journal of Accounting Research* 46:975-1016.

Bernard, V. and J. Thomas. 1990. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. *Journal of Accounting and Economics* 13:305-340.

Bhattacharya, U. and M. Krishnan. 1999. To believe or not to believe. *Journal of Financial Markets* 2:69-98.

Bodurtha Jr., J. and N. Mark. 1991. Testing the CAPM with time-varying risks and returns. *The Journal of Finance* 46:1485-1505.

Brennan, M., T. Chordia, A. Subrahmanyam, and Q. Tong. 2012. Sell-order liquidity and the cross-section of expected stock returns. *Journal of Financial Economics* 105:523-541.

Burgstahler, D. and I. Dichev. 1997. Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics* 24:99-126.

Bushee, B. and T. Goodman. 2007. Which institutional investors trade based on private information about earnings and returns? *Journal of Accounting Research* 45:289-321.

Caballé, J. and M. Krishnan. 1994. Imperfect competition in a multi-security market with risk neutrality. *Econometrica* 62:695-704.

Campbell, J., T. Ramadorai, and A. Schwartz. 2009. Caught on tape: Institutional trading, stock returns, and earnings announcements. *Journal of Financial Economics* 92:66-91.

Cho, Jin-Wan. 2007. Earnings announcements, private information, and strategic informed trading. *Journal of Financial Intermediation 16*:117-149.

Choe, H., B. Kho, and R. Stulz. 2005. Do domestic investors have an edge? The trading experience of foreign investors in Korea. *Review of Financial Studies* 18:795-829.

Collins, D., S. Kothari, and J. Rayburn. 1987. Firm size and the information content of prices with respect to earnings. *Journal of Accounting and Economics* 9:111-138.

Daley, L., J. Hughes, and J. Rayburn. 1995. The impact of earnings announcements on the permanent price effects of block trades. *Journal of Accounting Research* 33:317-334.

Davidson, R. and E. Flachaire. 2008. The wild bootstrap, tamed at last. *Journal of Econometrics* 146:162-169.

Dávila, E. and C. Parlatore. 2018. Identifying price informativeness. Working Paper, Yale University.

Duarte, J. and L. Young. 2009. Why is PIN priced? *Journal of Financial Economics* 91:119-138.

Dvořák, T. 2005. Do domestic investors have an informational advantage? Evidence from Indonesia. *Journal of Finance* 60:817-839.

Easley, D., N. Kiefer, M. O'Hara, and J. Paperman. 1996. Liquidity, information, and infrequently traded stocks. *Journal of Finance* 51:1405-1436.

Edelen, R., O. Inze, and G. Kadlec. 2016. Institutional investors and stock return anomalies. *Journal of Financial Economics* 119:472-488.

Fama, E. and K. French.2016. A five-factor asset pricing model. *Journal of Financial Economics* 116:1-22.

Foster, D. and S. Viswanathan. 1995. Can speculative trading explain the volume-volatility relation? *Journal of Business & Economic Statistics* 13:379-396.

Glosten, L. and P. Milgrom. 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics* 14:71-100.

Gompers, P. and A. Metrick. 2001. Institutional investors and equity prices. *Quarterly Journal of Economics* 116:229-259.

Gonedes, N. 1978. Corporate signaling, external accounting, and capital market equilibrium: Evidence on dividends, income, and extraordinary items. *Journal of Accounting Research* 16:26-79.

Griffin, J., T. Shu, and S. Topaloglu. 2012. Examining the dark side of financial markets: Do institutions trade on information from investment bank connections? *Review of Financial Studies* 25:2155-2188.

Grinblatt, M. and M. Keloharju. 2000. The investor behavior and performance of various investor types: A study of Finland's unique data set. *Journal of Financial Economics* 55:43-67.

Hansen, L. 1982. Large sample properties of generalized method of moments estimators. *Econometrica* 50:1029-1054.

Hayn, C. 1995. The information content of losses. *Journal of Accounting and Economics*. 20:125-153.

Hellwig, M. 1980. On the aggregation of information in competitive markets. *Journal of Economic Theory* 22:477-498.

Hendershott, T., D. Livdan, and N. Schürhoff. 2015. Are institutions informed about news? *Journal of Financial Economics* 117: 249-287.

Hirshleifer, D., S. Lim, and S. Teoh. 2009. Driven to distraction: Extraneous events and underreaction to earnings news. *Journal of Finance* 64:2289-2325.

Hu, G., B. Ke, and Y. Yu. 2018. Can transient institutions correctly interpret small negative earnings surprises in the absence of access to management's private information? *Journal of Accounting, Auditing & Finance* 33:3-33.

Huang, R. and C. Shiu. 2009. Local effects of foreign ownership in an emerging financial market: Evidence from qualified foreign institutional investors in Taiwan. *Financial Management* 38:567-202.

Jagannathan, R. and Z. Wang. 1996. The conditional CAPM and the cross-section of expected returns. *The Journal of Finance* 51:3-53.

Kang, J. and R. Stulz. 1997. Why is there a home bias? An analysis of foreign portfolio equity ownership in Japan. *Journal of Financial Economics* 46:3-28.

Kyle, A. 1985. Continuous auctions and insider trading. *Econometrica* 53:1315-1335.

Lundholm, R. 1988. Price-signal relations in the presence of correlated public and private information. *Journal of Accounting Research* 26:107-118.

Manzano, C. 1999. Price-signal relations in an imperfectly competitive financial market with public and private information. *Journal of Accounting Research* 37:451-463.

Nofsinger, J. and R. Sias. 1999. Herding and feedback trading by institutional and individual investors. *Journal of Finance* 54:2263-2295.

Puckett, A. and X. Yan. 2011. The interim trading skills of institutional investors. *Journal of Finance* 66:601-633.

Rochet, J. and J. Vila. 1994. Insider trading without normality. *The Review of Economic Studies* 61:131-152.

Shin, H. 1994. News management and the value of the firm. *The RAND Journal of Economics* 25:58-71.

Wood, S. 2001. Minimizing model fitting objectives that contain spurious local minima by bootstrap restarting. *Biometrics* 57:240-244.

Yan, X. and Z. Zhang. 2009. Institutional investors and equity returns: Are short-term institutions better informed? *Review of Financial Studies* 22:893-924.

Appendix A: Proof of Proposition 1

Given our assumption about component payoff structure, the multinormal random vector $\tilde{y} \equiv$ tr$\{\tilde{v}_F, \tilde{v}_T, \tilde{z}\}$, where "tr" denotes the transpose, is

$$\begin{bmatrix} \tilde{v}_F \\ \tilde{v}_T \\ \tilde{z} \end{bmatrix} \sim MN \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_F^2 & \rho.\sigma_F.\sigma_T & 0 \\ \rho.\sigma_F.\sigma_T & \sigma_T^2 & 0 \\ 0 & 0 & \sigma_z^2 \end{bmatrix} \right\}.$$

Let us call this 3x3 variance-covariance matrix Σ, and let tr(j) $\equiv \{1,1\}$. Let $\tilde{y}_1 \equiv$tr$\{\tilde{v}_F, \tilde{v}_T\}$, and Σ_{11} be the leading 2x2 minor of Σ. Then total payoff $\tilde{v} \equiv$ tr(j).$\tilde{y}_1$, and $\tilde{v} \sim N(0,$ tr(j).Σ_{11}.j).

We now define the strategic trader's optimization problem. Since the trader can observe announced earnings v_F, her own private information v_T, the noise trade z, and faces the pricing rule $p = \alpha + \beta v_F + \lambda\omega$, where aggregate order flow $\omega = x + z$, the problem of the trader is to choose a demand x to maximize profit, defined by $((v_F + v_T) - (\alpha + \beta v_F + \lambda(x + z)))x$ which yields the first-order condition $-\alpha + (1 - \beta)v_F + v_T - \lambda z = 2\lambda x$. Solving for x yields $\tau_0 = \frac{-\alpha}{2\lambda}$, $\tau_1 = \left(\frac{(1-\beta)}{2\lambda}\right)$, $\tau_2 = \left(\frac{1}{2\lambda}\right)$, $\tau_3 = -\left(\frac{1}{2}\right)$.

Then, aggregate order flow $\omega = x + z = \frac{-\alpha}{2\lambda} + \tau_1 v_F + \tau_2 v_T + \left(\frac{1}{2}\right)$z.

We then compute the expectation $E(\tilde{v}|v_F, \omega)$ where $\tilde{v} = \tilde{v}_F + \tilde{v}_T$. Define the multinormal random vector $\tilde{h} \equiv$ tr$\{\tilde{v}, \tilde{v}_F, \tilde{\omega}\}$, where "tr" denotes the transpose.

$$\begin{bmatrix} \tilde{v} \\ \tilde{v}_F \\ \tilde{\omega} \end{bmatrix} \sim MN \left\{ \begin{bmatrix} 0 \\ 0 \\ \frac{-\alpha}{2\lambda} \end{bmatrix}, \begin{bmatrix} \text{tr(j)}.\Sigma_{11}.\text{j} & \sigma_F^2 + \rho.\sigma_F.\sigma_T & 0 \\ \sigma_F^2 + \rho.\sigma_F.\sigma_T & \sigma_F^2 & 0 \\ 0 & 0 & \boldsymbol{Var}(\tilde{\omega}) \end{bmatrix} \right\},$$

where $Var(\tilde{\omega}) = Var(\tilde{x} + \tilde{z}) = Var(\tilde{x}) + Var(\tilde{z}) + 2Cov(\tilde{x}, \tilde{z})$.

Because of multinormality, the expectation $E(\tilde{v}|v_F, \omega)$is linear in the conditioning arguments. Recall that by virtue of market efficiency we have $p = E(v|v_F, \omega)$. Therefore, we equate corresponding coefficients to obtain three equations of the form, $\alpha = f_1(\alpha, \beta, \lambda)$, $\beta = f_2(\alpha, \beta, \lambda)$, $\lambda = f_3(\alpha, \beta, \lambda)$. From the first alone, it is easy to show that $\alpha = 0$. Manipulating the other two leads to a cubic in two variables, β and λ, instead of in λ alone as in Kyle (1985) and Rochet and Vila (1994). We obtain three candidate solutions of which only one satisfies $\lambda > 0$, which is needed to satisfy second-order conditions. So, we have a unique real root. The solution is easily verified. Plugging the equilibrium values of β and λ into the trader's strategy coefficients yields Proposition 1.

Appendix B: Variable Definitions

Variable	**Data Source**	**Explanation**
ERET	Prowess	Earnings announcement return obtained by compounding raw returns over days [0, 1] relative to the earnings announcement date.
EPS	Prowess	Basic Earnings per Share before Extraordinary Items
UE*	Prowess	EPS for quarter t less EPS for quarter t-4 divided by closing price at the beginning of quarter t.
FIITR*	SEBI	Net FII buying over the earnings announcement period, days [0, 1], divided by shares outstanding. Net FII buying for a firm on a day equals the number of shares bought, less the number of shares sold for that firm by all FIIs on that day.
MRET	Prowess	Return on the CNX Nifty Index compounded over days 0 and 1. The daily index return is calculated as the daily percentage change in the Index.
MCAP	Prowess	Market capitalization at the beginning of the quarter.
LMCAP	Prowess	Log of MCAP.
BM	Prowess	Book value of equity at the end of the most recent fiscal year before the earnings announcement (year t-1) divided by MCAP.
MOM3	Prowess	Three-month return during the fiscal quarter before the earnings announcement date.
OPROF	Prowess	Profit before interest, tax, and depreciation for year t-1 divided by total assets at the end of year t-2.
AGRO	Prowess	Percentage change in total assets in year t-1.
STDRET	Prowess	Standard deviation of daily returns over the fiscal quarter before the earnings announcement date.

VOL	Prowess	Monthly volume divided by shares outstanding, measured for the third month of the quarter before the fiscal quarter for which earnings is announced.
LAG_UE	Prowess	Value of UE lagged by one quarter.
L_AGE	Prowess	Logarithm of age of the firm in years at the end of quarter t relative to the year of incorporation.
DIVY	Prowess	Annual dividend in year t-1 divided by MCAP.
LPRC	Prowess	Logarithm of beginning quarter price.
CH_WK_EXCH	Prowess	Change in the US Dollar-Rupee rate over the week ending on day -1 (or day -29) relative to the earnings announcement period.
WK_MRET	Prowess	Market return over the week ending on day -1 (or day -29).
AFE	IBES	EPS for quarter t less mean analyst forecast per share for quarter t-4 divided by closing price at the beginning of quarter t.
Attention	Prowess	Average number of market-wide earnings announcements over days [0, 1]
NON_ERET	Prowess	Non-announcement period return obtained by compounding raw returns over days [-31, 30] relative to the earnings announcement date
NON_FIITR*	SEBI	Net FII buying over the non-announcement period, days [-31, 30], divided by shares outstanding.

* UE, FIITR, and NON_FIITR are centered and scaled in the manner described in Section 4.2

Appendix C: Summary of GMM and Bootstrap procedures for estimates and standard errors
Parameter Estimates

Our GMM approach to estimate the pricing model is to specify moment conditions of the form $E(u'Z) = 0$, where the instrument, Z, is plausibly orthogonal to the pricing error u. We chose five instruments that are in turn the constant (1), v_F (UE), ω (FIITR), the pre-announcement week change in the US Dollar-Rupee rate (CH_WK_EXCH), and the pre-announcement week market return (WK_MRET). The GMM objective function, Q(.), is $\overline{h(\Theta)}'W\overline{h(\Theta)}$ where $\overline{h(\Theta)}$ is the sample mean of the five moment conditions (5×1), and W is a positive definite and symmetric 5×5 matrix of weights. The objective function is nonlinear in the coefficients (though linear in the variables), and continuous and smooth. Because of the nonlinearity, we need to compute GMM estimates by numerically minimizing the objective function. If the criterion function were convex, then the objective function has a unique local minimum, which is also the global minimum. With convexity, an optimization program with any set of starting values should be able to reach a global minimum.

Since we have an ill-behaved objective function, there are many local minima. This was confirmed when we used an initial round of starting values. So, it became crucial to select good starting values. To aid in this, as a first step, we evaluated (*without any estimation*) the GMM objective function, assuming a weight matrix defined by the inverse of the moment condition variance-covariance matrix, at each of a dense grid of 7.96 million sets of parameter values. The 7.96 million sets of parameter values are obtained because we allow, ρ to vary between -0.99 and 0.99 in steps of 0.01 (total of 199 values), σ_1 to vary between 0.25 and 50 in steps of 0.25 (total of 200 values), and σ_2 to vary between 0.25 and 50 in steps of 0.25 (total of 200 values). This gives us a total of 199×200×200 = 7.96 Million.

While for the actual estimation using a gradient search, even a hundred starting points is quite computer-intensive, for merely evaluating the objective function, doing so even at 7.96 million points (with the optimal weight matrix) is quite feasible. This is because the computation is simplified by virtue of having to invert the relevant data matrices only once and needing to update only the coefficients β and λ for each set of starting points.[31]

We then plotted the objective function in turn against each combination of 2 of our 3 parameters. An inspection of the plots confirmed the ill-behavedness of the objective function. We also noticed some clustering of objective function values. From the lowest objective function values attained from the 7.96 million parameter value grid, we defined 100 "best" starting values for the three parameters. Note that even these best values would, in general, be along a gradient and not at a local minimum. This initial step enabled us to learn about how low an objective function value we could hope for when we subsequently implement GMM.

For each of the 100 starting value sets, we applied the classical iterated GMM algorithm (Hansen (1982)), modified to incorporate a perturbation step proposed by Wood (2001), as under. Let P_{k-1} be the initial parameter set in stage k; V_k, the objective function value in stage k, and let O and B denote

[31] Exactly how the covariance computation is simplified is explained in detail in Internet Appendix IA.B.

the original and (pairs) bootstrap sample at each stage, respectively. The Wood Perturbation is as follows:

1. Given P_{k-1} and the original sample, compute an optimal parameter set P_{kO1} yielding objective function value V_{kO1}.
2. Given P_{k-1} and a bootstrap sample, compute an optimal parameter set P_{kB}.
3. Given P_{kB} and the original sample, compute an optimal parameter set P_{kO2} yielding objective function value V_{kO2}.
4. $P_k = P_{kOl}$, $l \in \{1, 2\}$ and $l =$ argmin$\{V_{kO}\ , V_{kO2}\}$, and $V_k =$ min$\{V_{kO1}, V_{kO2}\}$.

Convergence is defined as $|P_{k-1} - P_k|$ and $|V_{k-1} - V_k|$ being within the tolerance limit, which we defined to equal 1e-8 (0.00000001). In the estimation, since the equilibrium holds only for strictly interior values of $\rho, \sigma_1,$ and σ_2, we define e = 0.0001, and use this to impose the following restrictions, $\rho \in$ [-0.9999, 0.9999]; $\sigma_1 \geq 0.0001$; $\sigma_2 \geq 0.0001$. We also imposed an upper limit on the number of iterations and the time allowed. Note that even when this upper limit is reached before convergence per the above definition, the estimates that we have at the last iteration are like enhancements to a 2-step GMM estimate, and so we also include them in our analysis, as in the firm-by-firm estimation when we encounter several such cases.

The Wood perturbation reduces the chance of being trapped in a flat region of the objective function and makes it more likely that we move to a better local minimum. The perturbation step is applied to each of the 100 sets of starting values to obtain 100 sets of parameter estimates. From these 100 sets of estimates obtained, we choose the final parameter estimates as the ones with the smallest objective function value. The local minimum thus achieved was smaller than the lowest minimum from the simple evaluation of the objective function on the dense initial grid.

For the firm-by-firm estimation discussed in section 6.3.2, to obtain good starting points, we evaluated each firm's objective function in the manner described in this Appendix, at one million points, and chose the ten best points as starting values.

Standard Errors

To obtain standard errors for the parameter estimates, we use a wild bootstrap procedure. To implement this procedure, we first compute model residuals based on the final parameter estimates. The residuals are then multiplied by an independent Rademacher random variable (+1 or -1, with equal probability). The use of the Rademacher transformation makes the bootstrap design "wild," and adjusts for heteroscedasticity (Davidson and Flachaire (2008)). The transformed residuals, final parameter estimates, and actual values of UE and FIITR are then used to define predicted ERET. We repeat this procedure 1,000 times to generate 1,000 bootstrap samples of predicted ERET. We then apply Iterated GMM to the predicted ERET, and the values of the five instruments to obtain 1,000 sets of bootstrap parameter estimates. The standard deviations of these parameter estimates serve as standard errors.

As an alternative, we also compute wild *cluster* bootstrap standard errors. Here, from the vector of the original residuals based on final parameter estimates, we define observation clusters by firm and resample entire clusters with replacement to define the residuals that will enter a given bootstrap sample. Bootstrapping from firm clusters attempts to account for serial correlation in the errors of each firm. Note that since we define observation clusters by firm, only the number of clusters is constant. Since the clusters vary in the number of observations, and we resample clusters randomly, the size of each bootstrap sample will, in general, be different. The Rademacher transformation is then applied to entire clusters. The remainder of the procedure to compute standard errors is identical to the procedure to obtain wild bootstrap standard errors.

For the results reported in Table 4, we find that the wild bootstrap standard errors are approximately 98% of the size of the wild cluster bootstrap standard errors. Since our inferences are very similar for the two procedures, for the remaining tables in the paper, we report results based on the simpler wild bootstrap.

$\boldsymbol{\sigma_F}$ and its Standard Error

One of the parameters of our pricing model is the variance of the private information of the firm, σ_F. In all the tables where we report GMM estimates of the primitive parameters, we equate σ_F to the standard deviation of unexpected earnings (UE). To compute the standard error of σ_F, we resample from the UE vector 1,000 times to construct 1,000 bootstrapped UE vectors. We then compute the standard deviation of each resampled UE vector. The standard deviation of these standard deviations is the standard error of σ_F.

Table 1

Descriptive Statistics

This table presents descriptive statistics for the variables used in the estimation of the pricing model obtained in Proposition 1 (regime 3 model). The sample consists of 17,877 earnings announcements for the years 2003 to 2016 for which net FII buying is non-zero on the earnings announcement date. Data on FII trades are obtained from the website: https://www.cdslindia.com/publications/FII/EquityDataFII.htm. Quarterly earnings announcement dates, earnings per share, stock prices, firm and market returns, annual financial data, industry codes, and quarterly FII ownership levels are obtained from the PROWESS database. Data on the change in the US Dollar-Rupee rate over the week ending on day -1 relative to the earnings announcement period (CH_WK_EXCH) is from the Federal Reserve website. Variable definitions are in Appendix B.

	# of obs.	Mean	Median	Std. Dev.	Minimum	Maximum
ERET	17,877	-0.09%	-0.32%	5.87%	-27.27%	25.00%
UE	17,877	-0.18%	0.10%	3.77%	-45.99%	24.59%
FIITR	17,877	0.00%	0.00%	0.21%	-1.33%	1.39%
MRET	17,877	0.04%	0.07%	2.51%	-15.70%	12.86%
LMCAP	17,877	10.24	10.16	1.64	5.78	14.70
BM	17,877	0.59	0.38	0.66	-0.03	5.37
MOM3	17,877	9.50%	5.06%	29.51%	-62.22%	188.61%
OPROF	17,877	18.99%	16.34%	12.29%	-2.11%	88.86%
AGRO	17,877	22.33%	15.32%	32.74%	-34.15%	358.89%
STDRET	17,877	2.64%	2.43%	1.03%	0.88%	8.10%
VOL	17,877	27.98%	13.18%	45.93%	0.58%	977.67%
LAG_UE	17,877	-0.28%	0.11%	3.86%	-35.79%	23.32%
L_AGE	17,877	2.59	2.77	0.83	0.00	4.58
DIVY	17,877	1.56%	1.10%	1.58%	0.00%	10.57%
LPRC	17,877	5.40	5.41	1.23	1.84	9.15
CH_WK_EXCH	17,877	0.18%	0.10%	1.02%	-3.93%	4.36%
WK_MRET	17,877	-0.10%	0.04%	3.66%	-19.34%	17.31%

Table 2

Pearson Correlations

This table presents Pearson correlations for the variables used in the estimation of the pricing model obtained in Proposition 1 (regime 3 model). The sample consists of 17,877 earnings announcements for the years 2003 to 2016 for which net FII buying is non-zero on the earnings announcement date. Data on FII trades are obtained from the website: https://www.cdslindia.com/publications/FII/EquityDataFII.htm. Quarterly earnings announcement dates, earnings per share, stock prices and firm and market returns, annual financial data, industry codes, and quarterly FII ownership levels are obtained from the PROWESS database. Correlations that are significant at the 1% level are marked with an asterisk, *. Variable definitions are in Appendix B.

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
(1) ERET	1.00														
(2) UE	0.06*	1.00													
(3) FIITR	0.09*	-0.00	1.00												
(4) MRET	0.30*	-0.00	0.01*	1.00											
(5) LMCAP	-0.03*	-0.02*	-0.02*	-0.06*	1.00										
(6) BM	-0.00	-0.04*	-0.01	0.01*	-0.10*	1.00									
(7) MOM3	0.00	0.04*	0.03*	-0.00	0.01	-0.07*	1.00								
(8) OPROF	0.00	-0.00	-0.00	-0.00	0.02*	-0.01	-0.00	1.00							
(9) AGRO	0.00	-0.00	-0.00	-0.00	0.00	-0.00	-0.00	0.96*	1.00						
(10) STDRET	0.01	0.01	-0.01*	0.05*	-0.50*	0.06*	0.22*	-0.01*	-0.00	1.00					
(11) VOL	-0.03*	-0.00	-0.03*	0.00	0.12*	-0.02*	0.15*	0.00	0.00	0.12*	1.00				
(12) LAG_UE	0.01*	0.29*	-0.01	-0.00	0.01	-0.06*	0.05*	-0.02*	-0.01	-0.01	0.00	1.00			
(13) L_AGE	-0.01	0.02*	0.01	-0.02*	0.05*	-0.06*	0.00	-0.02*	-0.02*	-0.09*	-0.11*	0.01*	1.00		
(14) DIVY	0.04*	-0.03*	-0.00	0.02*	0.00	0.10*	-0.10*	0.01*	-0.00	-0.14*	-0.07*	-0.02*	0.02*	1.00	
(15) LPRC	0.01	-0.01*	0.01	-0.04*	0.71*	-0.11*	0.10*	0.04*	0.01	-0.47*	0.10*	0.02*	0.08*	0.07*	1.00

Table 3

Earnings Announcement Return Regressions

The table reports regressions of ERET on UE, FIITR, and control variables. Columns (1) to (4) are based on OLS, and column (5) is based on 2SLS. Firm and year effects are included in the estimation but not reported to conserve space. Standard errors are clustered by firm. The sample consists of 17,877 earnings announcements for the years 2003 to 2016 for which net FII buying is non-zero on the earnings announcement date. Data on FII trades are obtained from the website: https://www.cdslindia.com/publications/FII/EquityDataFII.htm. Quarterly earnings announcement dates, earnings per share, stock prices, firm and market returns, annual financial data, industry codes, and quarterly FII ownership levels are obtained from the PROWESS database. Data on the change in the US Dollar-Rupee rate over the week ending on day -1 relative to the earnings announcement period (CH_WK_EXCH) is from the Federal Reserve website. Variable definitions are in Appendix B.

	(1)		(2)		(3)		(4)		(5)	
	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value
Intercept	0.017	0.01	0.014	0.02	0.018	0.00	0.015	0.01		
UE			9.037	0.00			9.161	0.00	9.437	0.00
FIITR					5.219	0.00	5.261	0.00	17.602	0.01
MRET	0.858	0.00	0.860	0.00	0.848	0.00	0.849	0.00	0.826	0.00
LMCAP	-0.034	0.00	-0.034	0.00	-0.031	0.00	-0.031	0.00	-0.024	0.00
BM	-0.008	0.02	-0.004	0.27	-0.008	0.03	-0.003	0.39	-0.001	0.84
MOM3	0.004	0.01	0.003	0.10	0.002	0.22	0.001	0.71	-0.004	0.15
OPROF	-0.004	0.09	-0.001	0.80	-0.006	0.02	-0.002	0.43	-0.005	0.13
AGRO	-0.003	0.10	-0.002	0.27	-0.003	0.07	-0.002	0.21	-0.003	0.14
STDRET	0.003	0.16	0.004	0.06	0.004	0.04	0.005	0.01	0.007	0.00
VOL	-0.013	0.00	-0.012	0.00	-0.013	0.00	-0.012	0.00	-0.013	0.00
LAGUE	0.005	0.00	-0.001	0.60	0.005	0.00	-0.001	0.45	-0.002	0.31
L_AGE	0.005	0.51	0.005	0.54	0.006	0.46	0.006	0.49	0.007	0.46
DIVY	0.000	0.91	0.003	0.30	0.000	0.94	0.002	0.39	0.001	0.76
L_PRC	0.001	0.73	0.001	0.84	0.000	0.97	0.000	0.91	-0.003	0.46
Firm and Year Effects		Yes		Yes		Yes		Yes		Yes
Number of Clusters		1,132		1,132		1,132		1,132		1,132
Number of Obs.		17,877		17,877		17,877		17,877		17,877
Adjusted R^2		14.88%		16.77%		18.13%		20.08%		18.24%
Durbin-Wu-Hausman Test of Endogeneity (p-value):										4.06 (0.04)
Weak Instrument Test (Kleibergen-Paap rank Wald F statistic):										10.99
Stock-Yogo Weak ID test critical values: 10% maximal LIML size:										8.68
Hansen J statistic (p-value):										1.42 (0.23)

Table 4

GMM Estimates of Primitive Parameters and Shallow Parameters

The table reports GMM estimates of the primitive and shallow parameters of the following model:
$p = (\alpha + \beta v_F + \lambda\omega)$, with $\alpha = 0, \beta = 1 + \rho \times \sqrt{\sigma_1^2}, \lambda = \sqrt{\sigma_2^2} \times \sqrt{1-\rho^2}$;
$\sigma_1 = \left(\frac{\sigma_T}{\sigma_F}\right)$ *and* $\sigma_2 = \left(\frac{\sigma_T}{\sigma_Z}\right)$.

To compute the three parameter estimates ρ, σ_1, and σ_2, we use the Iterated GMM Procedure of Hansen (1982) that is modified by a perturbation first proposed in Wood (2001). Standard errors are computed using a wild bootstrap procedure. Both the parameter estimation method and standard error computation are described in Appendix C. With an independently estimated σ_F, and estimates of σ_1 and $\sigma_2, \sigma_T = \sigma_F \times \sigma_1$, and $\sigma_Z = \left(\frac{\sigma_T}{\sigma_2}\right)$. We calculate σ_F as the sample standard deviation of unexpected earnings (UE). To compute its standard error, we resample from the UE vector 1,000 times to construct 1000 bootstrapped UE vectors. We then compute the standard deviation of each resampled UE vector. The standard deviation of these standard deviations is the standard error of σ_F.

p is the abnormal return compounded over the day of the earnings announcement and the following day, (0,1), v_F is the firm's unexpected earnings (UE), ω is the net FII buying over the earnings announcement period (FIITR). To implement GMM, we define moment restrictions of the form $E(u'Z) = 0$, where u is the pricing error, and the five instruments Z are in turn the constant (1), UE, FIITR, the change in the US Dollar-Rupee rate over the week ending on day -1 relative to the earnings announcement period (CH_WK_EXCH), and the market return over the week ending on day -1 (WK_MRET). Data sources are described above Table 1 and variable definitions are in Appendix B. In column (1), we report results based on the regime 3 sample that consists of earnings announcements (days [0, 1]) when both UE and FIITR are present. Column (2) contains the results when the regime 3 sample is augmented with earnings announcements by firms in that sample, for which there was no FII trading over days [0, 1].

	(1)		(2)	
	Regime 3 Sample		Regime 3 Sample and "No Trades"	
Parameter	Estimate	p-value	Estimate	p-value
ρ	0.249	0.00	0.282	0.00
σ_1	23.750	0.00	7.750	0.00
σ_2	4.724	0.00	5.749	0.00
σ_F	0.001	0.00	0.002	0.00
σ_T	0.023	0.00	0.018	0.00
σ_Z	0.005	0.00	0.003	0.00
β	6.908	0.00	3.183	0.00
λ	4.575	0.00	5.516	0.00
# of Obs.		17,877		39,346
Hansen J-Statistic (p-value)		0.0002 (0.99)		0.0005 (0.99)

Table 5
Forecast Accuracy of GMM and OLS Models – Cross-Validation Exercise

This table reports on the forecast accuracy of the OLS and GMM models described in Tables 3 and 4, respectively. We implement a ten-fold cross-validation exercise that has the following steps:

1. Randomly divide the regime 3 sample (17,877 obs.) into ten groups.
2. Treat one group as the holdout or test dataset (1,787 obs.).
3. Treat the remaining groups as the training dataset (16,090 obs.).
4. Estimate the model on the training set and use the parameters obtained to predict the dependent variable (ERET) for the test dataset.
5. Compute forecast errors as the difference between realized ERET and predicted ERET for the test dataset and then compute the Root Mean Squared Error (RMSE) and the Mean Absolute Error (MAE).
6. Repeat steps 2 to 5 for each of the remaining 9 groups.

For the ten test sub-samples, columns (1)-(4) presents the RMSE, MAE, p-value from a t-test that the Mean forecast error equals 0, and the p-value from a Wilcoxon sign rank test that the median forecast error equals zero, respectively.

	OLS				GMM			
Test sample	RMSE	MAE	t-test p-value	Wilcoxon test p-value	RMSE	MAE	t-test p-value	Wilcoxon test p-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1	0.048	0.036	0.76	0.040	0.048	0.036	0.76	0.040
2	0.048	0.037	0.81	0.149	0.048	0.037	0.81	0.149
3	0.050	0.038	0.43	0.262	0.050	0.038	0.43	0.262
4	0.049	0.037	0.68	0.244	0.049	0.037	0.68	0.244
5	0.051	0.038	0.79	0.334	0.051	0.038	0.79	0.334
6	0.052	0.039	0.37	0.012	0.052	0.039	0.37	0.012
7	0.049	0.037	0.69	0.025	0.049	0.037	0.69	0.025
8	0.049	0.037	0.41	0.016	0.049	0.037	0.41	0.016
9	0.050	0.038	0.93	0.086	0.050	0.038	0.93	0.086
10	0.049	0.037	0.46	0.061	0.049	0.037	0.46	0.061

Table 6

Firm-by-firm GMM Estimates of Primitive Parameters and Shallow Parameters

The table reports the distribution of the firm-by-firm estimates of the primitive and shallow parameters of the following model:

$p = (\alpha + \beta v_F + \lambda\omega)$, with $\alpha = 0, \beta = 1 + \rho \times \sqrt{\sigma_1^2}, \lambda = \sqrt{\sigma_2^2} \times \sqrt{1 - \rho^2}$;

$\sigma_1 = \left(\frac{\sigma_T}{\sigma_F}\right)$ and $\sigma_2 = \left(\frac{\sigma_T}{\sigma_Z}\right)$.

With an independently estimated σ_F, and estimates of σ_1 and $\sigma_2, \sigma_T = \sigma_F \times \sigma_1$, and $\sigma_Z = \left(\frac{\sigma_T}{\sigma_2}\right)$.

The model is estimated for the 365 firms in the sample with at least twenty quarterly observations during the sample period. These 365 firms span 12,118 firm-quarters. Columns (1) – (7) contain the distribution of estimates for the 365 firms, and the last column (8) contains the estimates for the 12,118 firm-quarter sample. To compute the three parameter estimates ρ, σ_1, and σ_2, we use the Iterated GMM Procedure of Hansen (1982) that is modified by a perturbation first proposed in Wood (2001). Since the equilibrium holds only for strictly interior values of ρ, σ_1, and σ_2, we define ε = 0.0001, and use this to impose the following restrictions, $\rho \in [-0.9999, 0.9999]$; $\sigma_1 \geq 0.0001$; $\sigma_2 \geq 0.0001$. The parameter estimation method is described in Appendix C. We calculate σ_F as the sample standard deviation of earnings (UE). To compute its standard error, we resample from the UE vector 1,000 times to construct 1000 bootstrapped UE vectors. We then compute the standard deviation of each resampled UE vector. The standard deviation of these standard deviations is the standard error of σ_F.

p is the abnormal return compounded over the day of the earnings announcement and the following day, (0,1), v_F is the firm's unexpected earnings (UE), ω is the net FII buying over the earnings announcement period (FIITR). To implement GMM, we define moment restrictions of the form $E(u'Z) = 0$, where u is the pricing error, and the five instruments Z are in turn the constant (1), UE, FIITR, the change in the US Dollar-Rupee rate over the week ending on day -1 relative to the earnings announcement period (CH_WK_EXCH), and the market return over the week ending on day -1 (WK_MRET). Data sources are described above Table 1 and variable definitions are in Appendix B.

	# of firms (1)	Mean (2)	Median (3)	Std. Dev. (4)	Min (5)	Max (6)	t-stat (7)	All Firms (8)
ρ	365	0.321	0.795	0.786	-0.999	0.999	7.802	0.773
σ_1	365	29.659	22.049	27.282	0.000	182.378	20.770	8.072
σ_2	365	22.381	16.913	20.266	0.000	132.183	21.099	8.775
σ_F	365	0.001	0.000	0.001	0.000	0.004	19.240	0.001
σ_T	365	0.013	0.008	0.013	0.000	0.095	18.094	0.007
σ_Z	365	13.231	0.001	65.891	0.000	783.525	3.836	0.001
β	365	13.666	9.036	33.655	-99.892	178.493	7.758	7.241
λ	365	8.343	6.188	9.865	0.000	66.707	16.157	5.566
J-Stat	365	0.119	0.086	0.102	0.001	0.692		0.000
p-value	365	0.944	0.958	0.046	0.708	0.999		0.999
Obs. per firm	365	33.200	32.000	9.596	20.000	53.000		-

Table 7

Parameter Estimates for Sub-samples sorted on Size, Earnings per Share, and Investor Attention

The table reports primitive and shallow parameter estimates of the following model for sub-samples formed by sorting firms on (i) size (ii) the sign and size of earnings per share and (iii) investor attention

$p = (\alpha + \beta v_F + \lambda\omega)$, with $\alpha = 0, \beta = 1 + \rho \times \sqrt{\sigma_1^2}, \lambda = \sqrt{\sigma_2^2} \times \sqrt{1-\rho^2}$;

$\sigma_1 = \left(\frac{\sigma_T}{\sigma_F}\right)$ and $\sigma_2 = \left(\frac{\sigma_T}{\sigma_Z}\right)$.

With an independently estimated σ_F, and estimates of σ_1 and $\sigma_2, \sigma_T = \sigma_F \times \sigma_1$, and $\sigma_Z = \left(\frac{\sigma_T}{\sigma_2}\right)$.

Panel A reports parameter estimates based on size partitions. We measure size as the log of market capitalization at the beginning of the quarter (LMCAP). To form size groups, we divide the sample into ten deciles based on LMCAP. Firm-quarters in smallest three deciles are small firms, those in the next four deciles are medium firms, and those in the largest three deciles are large firms. In panel B, firms are classified based on the size and sign of earnings per share (EPS). Small-profit firms are defined as firms with quarterly EPS that is > ₹0.00 and ≤ ₹0.01; loss firms are firms with EPS < 0; and large-profit firms are firms with EPS > ₹0.01. In panel C, firms are divided into three groups based on mean investor attention. Attention is measured using the average number of market-wide earnings announcements over days [0, 1] and firms are divided into three groups – bottom 30% (high attention), middle 40% (medium attention), and top 30% (low attention).

To compute the three parameter estimates ρ, σ_1, and σ_2, we use the Iterated GMM Procedure of Hansen (1982) that is modified by a perturbation first proposed in Wood (2001). The parameter estimation method is described in Appendix C. We calculate σ_F as the sample standard deviation of earnings (UE).

p is the abnormal return compounded over the day of the earnings announcement and the following day, (0,1), v_F is the firm's unexpected earnings (UE), ω is the net FII buying over the earnings announcement period (FIITR). To implement GMM, we define moment restrictions of the form $E(u'Z) = 0$, where u is the pricing error, and the five instruments Z are in turn the constant (1), UE, FIITR, the change in the US Dollar-Rupee rate over the week ending on day -1 relative to the earnings announcement period (CH_WK_EXCH), and the market return over the week ending on day -1 (WK_MRET). Data sources are described above Table 1 and variable definitions are in Appendix B.

Panel A: Firm Size

	Small Firms (1)		Medium Firms (2)		Large Firms (3)	
Parameter	Estimate	p-value	Estimate	p-value	Estimate	p-value
ρ	0.511	0.00	0.379	0.00	0.109	0.00
σ_1	10.500	0.00	20.746	0.00	44.500	0.00
σ_2	3.249	0.00	4.759	0.00	8.749	0.00
σ_F	0.0014	0.00	0.0007	0.00	0.0005	0.00
σ_T	0.015	0.00	0.015	0.00	0.024	0.00
σ_Z	0.005	0.00	0.003	0.00	0.003	0.00
β	6.361	0.00	8.862	0.00	5.857	0.00
λ	2.794	0.00	4.404	0.00	8.697	0.00
# of Obs.		5,370		7,150		5,357
Hansen J-Statistic		0.0054		0.0002		0.0035
(p-value)		(0.997)		(0.999)		(0.998)

Panel B: Small Profit-Loss-Large Profit

	Small-Profit Firms (1)		Loss Firms (2)		Profit Firms (3)	
Parameter	Estimate	p-value	Estimate	p-value	Estimate	p-value
ρ	-0.758	0.01	0.902	0.00	0.354	0.00
σ_1	2.770	0.30	4.368	0.00	24.499	0.00
σ_2	10.178	0.00	6.454	0.00	4.979	0.00
σ_F	0.001	0.00	0.002	0.00	0.001	0.00
σ_T	0.003	0.30	0.010	0.00	0.017	0.00
σ_Z	0.000	0.32	0.001	0.00	0.003	0.00
β	-1.100	0.60	4.941	0.00	9.665	0.00
λ	6.635	0.00	2.784	0.00	4.657	0.00
# of Obs.		304		1,504		16,069
Hansen J-Statistic		0.0377		0.0446		0.0001
(p-value)		(0.981)		(0.978)		(0.999)

Panel C: Attention, Average Number of Simultaneous Earnings Announcements

	High Attention (1)		Medium Attention (2)		Low Attention (3)	
Parameter	Estimate	p-value	Estimate	p-value	Estimate	p-value
ρ	0.27	0.00	0.45	0.00	0.53	0.00
σ_1	19.25	0.00	15.75	0.00	10.74	0.00
σ_2	4.76	0.00	5.31	0.00	5.04	0.00
σ_F	0.001	0.00	0.001	0.00	0.001	0.00
σ_T	0.02	0.00	0.01	0.00	0.01	0.00
σ_Z	0.003	0.00	0.003	0.00	0.002	0.00
β	6.14	0.00	8.16	0.00	6.72	0.00
λ	4.59	0.00	4.73	0.00	4.27	0.00
# of Obs.		5,302		7,096		5,351
Hansen J-Statistic		0.0016		0.0006		0.0004
(p-value)		(0.999)		(0.999)		(0.999)

Table 8

Using Analyst Forecasts as Earnings Expectations

The table reports GMM estimates of the primitive and shallow parameters of the following model for two alternate measures of earnings expectations – Four Quarter Lagged Earnings per Share and Mean Analysts Forecast of Earnings per share used to compute v_F (unexpected earnings):

$p = (\alpha + \beta v_F + \lambda\omega)$, with $\alpha = 0, \beta = 1 + \rho \times \sqrt{\sigma_1^2}, \lambda = \sqrt{\sigma_2^2} \times \sqrt{1-\rho^2}$;

$\sigma_1 = \left(\frac{\sigma_T}{\sigma_F}\right)$ and $\sigma_2 = \left(\frac{\sigma_T}{\sigma_Z}\right)$.

With an independently estimated σ_F, and estimates of σ_1 and σ_2, $\sigma_T = \sigma_F \times \sigma_1$, and $\sigma_Z = \left(\frac{\sigma_T}{\sigma_2}\right)$.

Table 4 contains definitions of p, v_F, and ω. Both the parameter estimation method and standard error computation are described in Appendix C. Data sources are described above Table 1, and variable definitions are in Appendix B. In column (1), we report results based on unexpected earnings using Four Quarter Lagged Earnings per Share as Earnings Expectations (UE). Column (2) reports results based on unexpected earnings using Mean Analysts Forecast of Earnings per share as Earnings Expectations (AFE).

For the two models, we calculate σ_F as the sample standard deviation of UE and AFE, respectively. To compute their standard errors, we resample from the UE (AFE) vector 1,000 times to construct 1000 bootstrapped UE (AFE) vectors. We then compute the standard deviation of each resampled vector. The standard deviation of these standard deviations is the standard error of σ_F.

	(1)		(2)	
	Earnings Expectations = Lag-4 Quarterly EPS		Earnings Expectations = Mean Analyst Forecast EPS	
Parameter	Estimate	p-value	Estimate	p-value
ρ	0.523	0.00	0.981	0.00
σ_1	8.662	0.00	49.629	0.00
σ_2	9.595	0.00	40.769	0.00
σ_F	0.001	0.00	0.0001	0.00
σ_T	0.007	0.00	0.007	0.00
σ_Z	0.0007	0.00	0.0002	0.00
β	5.530	0.00	49.682	0.00
λ	8.179	0.00	7.923	0.00
# of Obs.		910		910
Hansen J-Statistic (p-value)		0.005 (0.99)		0.005 (0.99)

Table 9

Effect of Incorporating Non-Announcement Period Information

In this table, we jointly estimate the following two models:

$$u_1 = ERET - (\beta^{FT} * UE + \lambda^{FT} * FIITR) \quad (2)$$

with $\beta^{FT} = 1 + \rho * \sqrt{\sigma_1^2}$, $\lambda^{FT} = \sqrt{\sigma_2^2} * \sqrt{1-\rho^2}$, where $\sigma_1 = (\sigma_{T1}/\sigma_F)$ and $\sigma_2 = (\sigma_{T1}/\sigma_{Z1})$.

$$u_2 = NON\_ERET - (\lambda^T * NON\_FIITR) \quad (3)$$

where $\lambda^T = \sigma_3 = (\sigma_{T2}/\sigma_{Z2})$

In Eq. (2), ERET is the abnormal return compounded over the day of the earnings announcement and the following day, (0,1), UE is the firm's unexpected earnings, FIITR is the net FII buying over the earnings announcement period. In Eq. (3), NON_ERET and NON_FIITR are measured over days [-31, -30] relative to the earnings announcement date. Data sources are described above Table 1 and variable definitions are in Appendix B.

To implement GMM, we define two sets of moment restrictions of the form $E(u_i Z) = 0$, i = 1, 2 where u_i is the pricing error. For Eq. (2), the pricing error is u_1 and the five instruments are in turn the constant (1), UE, FIITR, the change in the US Dollar-Rupee rate over the week ending on day -1 relative to the earnings announcement period, and the market return over the week ending on day -1. For Eq. (2), the pricing error is u_2 and the instruments are: 1 (constant), NON_FIITR, the change in the US Dollar-Rupee rate over the week ending on day -29 relative to the earnings announcement period, and the market return over the week ending on day -29. The parameter estimation method and the calculation of standard errors are described in Appendix C.

Column 1 reproduces the deep and shallow parameter estimates from the baseline specification based only on Regime 3 (see Table 4). The next two columns, we impose a cross-regime restriction to identify the deep parameters in Eq. (3). In the first specification, we require $\sigma_T = \sigma_{T1} = \sigma_{T2}$ and estimate the four parameters σ_T, σ_{Z1}, σ_{Z2}, and ρ (the common σ_T specification) and in the second specification, we require $\sigma_Z = \sigma_{Z1} = \sigma_{Z2}$ and estimate the four parameters $\sigma_{T1}, \sigma_{T2}, \sigma_Z$, and ρ (the common σ_Z specification). Internet Appendix IA.A summarizes the specification of the pricing errors, u_1 and u_2, instruments employed to define moment conditions, estimable parameters, and parameter restrictions.

	Baseline Model		Common σ_T Model		Common σ_Z Model	
	(1)		(2)		(3)	
Parameter Estimate	Estimate	p-value	Estimate	p-value	Estimate	p-value
ρ	0.249	0.00	0.290	0.00	0.169	0.00
σ_1	23.750	0.00	20.484	0.00	35.123	0.00
σ_2	4.724	0.00	4.692	0.00	4.832	0.00
σ_F	0.001	0.00	0.001	0.00	0.001	0.00
σ_T	0.023	0.00	0.020	0.00		
σ_Z	0.005	0.00			0.007	0.00
σ_{T1}					0.034	0.00
σ_{T2}					0.032	0.00
σ_{Z1}			0.004	0.00		
σ_{Z2}			0.004	0.00		
β^{FT}	6.908	0.00	6.940	0.00	6.938	0.00
λ^{FT}	4.575	0.00	4.625	0.00	4.624	0.00
λ^{T}			4.428	0.00	4.428	0.00
# of Obs.		17,877		17,765		17,765
Hansen J-Statistic		0.0002		0.006		0.006
(p-value)		(0.99)		(0.99)		(0.99)

INTERNET APPENDIX

The Pricing of Earnings in the Presence of Informed Trades: A Simple GMM Approach

Figure IA.1

FII Buying and Selling Around Earnings Announcements

In this Figure, we plot median FII buying, FII selling, and *net* FII buying around earnings announcements. FII buying (selling) on a day aggregates all buys (sells) for that day and divides that sum by shares outstanding. Net FII buying for each firm-day equals the number of shares bought less the number of shares sold by all FIIs for that firm on that day, divided by shares outstanding. The sample consists of only those firm-quarters for which FII trading during the earnings announcement period is non-zero. Data on FII trading are obtained from the SEBI website: http://www.sebi.gov.in. Earnings announcement dates are from the PROWESS database.

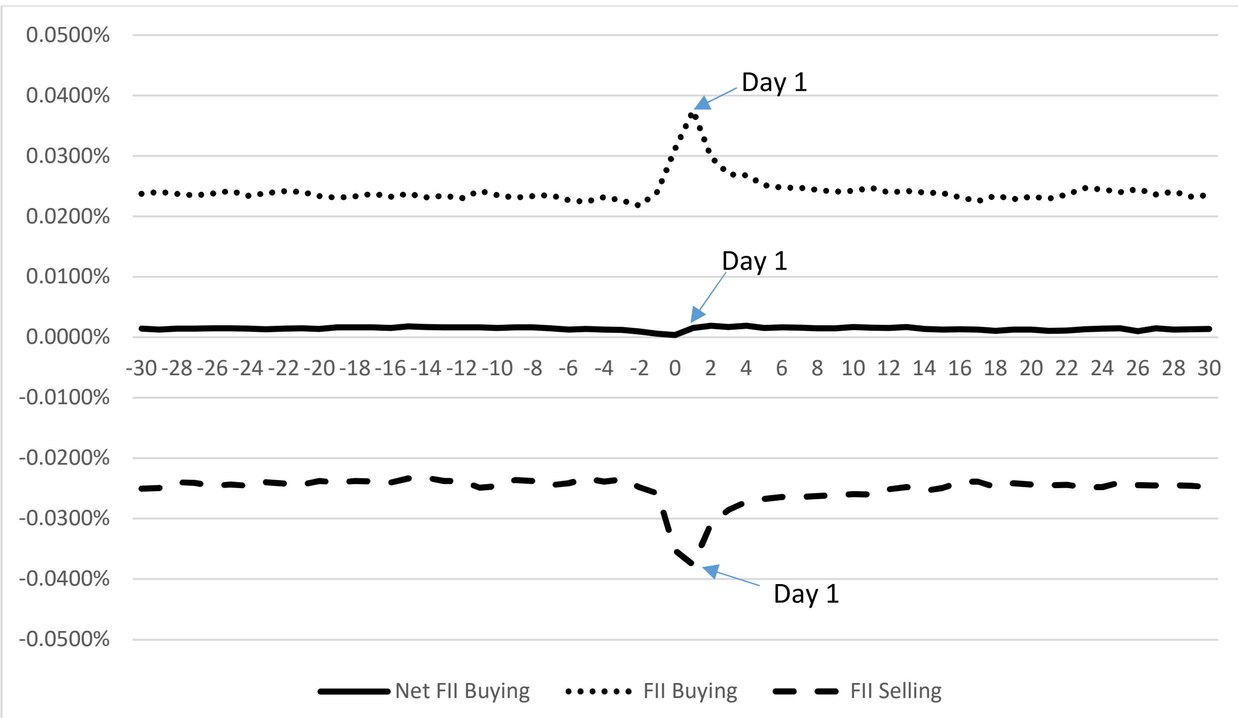

Figure IA.2

Histogram of Basic Earnings per share around zero

In this Figure, we present a partial histogram of earnings per share frequencies around zero, ranging from -2.0 ₹ per share to 2.0₹ per share. The sample consists of listed Indian firms for the years 2003-2016. Data on Earnings per share are from Prowess.

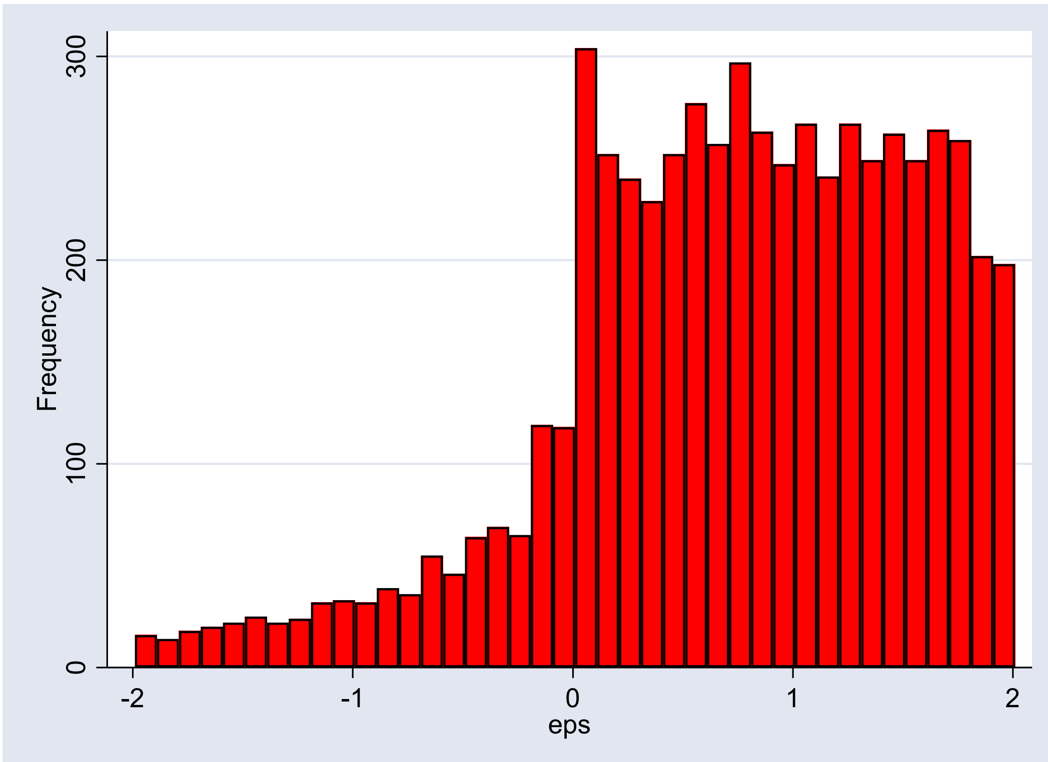

Table IA.1

Sample Selection

Our initial sample consists of all Indian firms listed on the National Stock Exchange (NSE) with non-missing quarterly earnings announcement dates and non-missing earnings per share for the years 2003-2016. Data on FII trades are obtained from the website: https://www.cdslindia.com/publications/FII/EquityDataFII.htm. Quarterly earnings announcement dates and earnings per share, stock prices and returns, annual financial data, industry codes, and quarterly FII ownership levels are obtained from the PROWESS database.

Initial Sample of Firm-quarters (2003-2016)	92,714
Less: Firm-quarters with missing data for unexpected earnings and its four-quarter lagged value	19,183
Less: Firm-quarters other than March, June, September, and December	78
Less: Firm quarters with erroneous earnings announcement dates or dates that are more than 180 days after the fiscal quarter-end	164
Less: Firm-quarters with more than 45 missing returns during the 90 trading-day period centered on the earnings announcement date	3,314
Less: Firm-quarters with missing returns on the earnings announcement days [0, 1]	1,750
Less: Firm-quarters with missing data on control variables	7,975
Less: Singleton firm-quarters	194
Final Sample	60,056
Composition of Final Sample:	
Firm-quarters with FII trading during earnings announcements (30%)	17,877
Firm-quarters with no FII trading during earnings announcements and with FII ownership (40%)	24,101
Firm-quarters with no FII trading during earnings announcements and with *no* FII ownership (30%)	18,078

Table IA.2

Yearly Distribution of FII Trading during Earnings Announcements

This table reports the sample distribution by year for three types of firm-quarters: (a) firm-quarters with FII trading during earnings announcements and (b) firm-quarters with no FII trading during earnings announcements, when FIIs own shares, and (c) Zero FII Ownership. The sample period consists of the years 2003 to 2016. Data on FII trades are obtained from the website: https://www.cdslindia.com/publications/FII/EquityDataFII.htm. Quarterly earnings announcement dates and earnings per share, stock prices and returns, annual financial data, industry codes, and quarterly FII ownership levels are obtained from the PROWESS database.

	Trading During Earnings Announcements							
	Non-Zero		Zero		Zero FII ownership		Total	
Year	Num.	%	Num.	%	Num.	%	Num.	%
2003	384	14%	1,223	44%	1,186	42%	2,793	100%
2004	543	19%	1,158	41%	1,105	39%	2,806	100%
2005	1,098	28%	1,574	40%	1,294	33%	3,966	100%
2006	1,310	31%	1,710	41%	1,185	28%	4,205	100%
2007	1,426	32%	1,728	39%	1,240	28%	4,394	100%
2008	1,034	31%	1,394	42%	905	27%	3,333	100%
2009	1,154	28%	1,747	43%	1,153	28%	4,054	100%
2010	1,382	33%	1,625	38%	1,221	29%	4,228	100%
2011	1,411	33%	1,705	39%	1,221	28%	4,337	100%
2012	1,406	29%	1,969	41%	1,436	30%	4,811	100%
2013	1,584	31%	2,116	42%	1,366	27%	5,066	100%
2014	1,726	30%	2,357	41%	1,734	30%	5,817	100%
2015	1,921	33%	2,281	39%	1,679	29%	5,881	100%
2016	1,498	34%	1,514	35%	1,353	31%	4,365	100%
Total	17,877	30%	24,101	40%	18,078	30%	60,056	100%

Appendix IA.A: Summary of GMM Specifications

Panel A: Pricing errors and available instruments for each regime.

Regime	Pricing error u	Instruments Z
3	$u_1 = p^{FT} - (\beta^{FT} * v_F + \lambda^{FT} * \omega^{FT})$ $= ERET - (\beta^{FT} * UE + \lambda^{FT} * FIITR)$ with $\beta^{FT} = 1 + \rho * \sqrt{\sigma_1^2}$, $\lambda^{FT} = \sqrt{\sigma_2^2} * \sqrt{1-\rho^2}$ where $\sigma_1 = (\sigma_{T1}/\sigma_F)$ and $\sigma_2 = (\sigma_{T1}/\sigma_{Z1})$.	5 variables: 1 (constant), UE, FIITR, CH_USD_INR, L_MRET (measured during or relative to the earnings announcement period, days [0, 1].
2	$u_2 = p^T - (\alpha^T + \lambda^T * \omega^T)$ $= NON\_ERET - (\alpha^T + \lambda^T * NON\_FIITR)$ with $\lambda^T = \sqrt{\sigma_3^2}$.	4 variables: 1 (constant), NON_FIITR, CH_USD_INR, L_MRET, (measured during or relative to the non-announcement period, days [-31, -30].

All moment conditions used are of the form $E(u_i * Z) = 0, i = 1,2$ where u is a pricing error and Z is one data variable. In all cases σ_F is estimated independently, before the GMM estimation of other primitive parameters. ERET, NON_ERET, UE, FIITR, NON_FIITR, CH_USD_INR, and L_MRET are defined in Appendix B.

Panel B: List of Possible GMM specifications with the number of moment conditions, and the list of parameters and parameter restrictions

	Conditions and data from	# conditions	Parameters	Parameter restrictions
1	Regime 3 alone	5	σ_1, σ_2, ρ	$\sigma_1 > 0, \sigma_2 > 0, \rho \in (-1,1)$
2	Regimes 3 and 2	9	$\sigma_1, \sigma_2, \sigma_3, \rho$	$\sigma_1 = (\sigma_{T1}/\sigma_F), \sigma_2 = (\sigma_{T1}/\sigma_{Z1}), \sigma_T > 0, \sigma_Z > 0, \sigma_3 > 0, \rho \in (-1,1)$
3	Regimes 3 and 2 (σ_T constant)	9	$\sigma_T, \sigma_{Z1}, \sigma_{Z2}, \rho$	$\sigma_1 = (\sigma_{T1}/\sigma_F), \sigma_2 = (\sigma_{T1}/\sigma_{Z1}), \sigma_3 = (\sigma_{T2}/\sigma_{Z2}); \sigma_{T1} = \sigma_{T2}, \sigma_{Z1} > 0, \sigma_{Z2} > 0, \rho \in (-1,1)$
4	Regimes 3 and 2 (σ_Z constant)	9	$\sigma_{T1}, \sigma_{T2}, \sigma_Z, \rho$	$\sigma_1 = (\sigma_{T1}/\sigma_F), \sigma_2 = (\sigma_{T1}/\sigma_{Z1}), \sigma_3 = (\sigma_{T2}/\sigma_{Z2}); \sigma_{T1} > 0, \sigma_{T2} > 0, \sigma_{Z1} = \sigma_{Z2}, \rho \in (-1,1)$
5	Regimes 3 and 2	9	$\sigma_{T1}, \sigma_{T2}\ \sigma_{Z1}, \sigma_{Z2}, \rho$	$\sigma_1 = (\sigma_{T1}/\sigma_F), \sigma_2 = (\sigma_{T1}/\sigma_{Z1}), \sigma_3 = (\sigma_{T2}/\sigma_{Z2}); \sigma_{T1} > 0, \sigma_{T2} > 0, \sigma_{Z1} > 0, \sigma_{Z2} > 0, \rho \in (-1,1)$

Under specification 2, for Regime 2, where $\sigma_3 = \lambda^T$, and $\lambda^T = (\sigma_{T2}/\sigma_{Z2})$, we cannot solve for σ_{T2} and σ_{Z2}. In specification 5, the solutions for σ_{T2} and σ_{Z2} are valid only up to a scalar multiple. Under each of the other two-regime specifications, we can identify all the primitive parameters in each regime.

Appendix IA.B: Details of Covariance Matrix Calculation To Define Starting Values

This Appendix should be read together with Internet Appendix D, which provides a summary of all the models and moment conditions that we use in our estimation. A key step in our GMM approach is in the selection of good starting values by simply evaluating the GMM objective function, in most cases, at 7.96 million points. (For the firm-by-firm estimation, to obtain good starting points we evaluate each firm's objective function in the manner described in this Appendix, at one million points, and chose the ten best points as starting values.) We use the optimal weight matrix at each point which is the inverse of the covariance matrix of the moment conditions. It helps a lot to do this evaluation in a less computer-intensive way. We provide details below that will help a reader more easily understand the R code that implements the estimation.

<u>For the baseline specification (results tabulated in Tables 4 to 8)</u>

Suppressing the superscript "(FT)" which denotes Regime 3 which has both earnings and trading signals, we use the following notation (with "T" without the parentheses denoting the matrix transpose) to define the moment conditions more concisely:

i) Instrument vector $\tilde{z}^T = [\tilde{z}_1, \tilde{z}_2, \tilde{z}_3, \tilde{z}_4, \tilde{z}_5]$, where $\tilde{z}_1 = 1$, $\tilde{z}_2 = \tilde{v}_F$, $\tilde{z}_3 = \widetilde{\omega}$, $\tilde{z}_4 =$ CH_WK_EXCH and $\tilde{z}_5 =$ WK_MRET. Data definitions are in Appendix B.
ii) random vector listing variables that help define the pricing error $\tilde{x}^T = [\tilde{p}, \tilde{v}_F, \widetilde{\omega}]$,
iii) the 5-by-3 random matrix $\tilde{T} \equiv \tilde{z}.\tilde{x}^T$and
iv) parameters $B^T = [1, -\beta, -\lambda]$, where the shallow parameters β and λ are functions of the primitive parameters σ_F, σ_T, σ_z, and ρ, as defined in Proposition 1.

The five moment conditions in our primary specification can then be written as $E(\tilde{T}.B) = 0$. The optimal weight matrix in the GMM objective function is given by the inverse of the 5-by-5 covariance matrix of these five moment conditions. In our initial dense grid evaluation of the GMM objective function at 7.96 million points, efficiently computing this covariance matrix is important, as it can be an extremely computer-intensive step.

To do so, first, define each 3-element row vector of the 5-by-3 matrix $\tilde{T}$ as $\tilde{t}_k, k = 1,2,3,4,5$. Then the typical (i,j) element of the 5-by-5 covariance matrix of $\tilde{T}.B$ is given by $B^T.Cov(\tilde{t}_i, \tilde{t}_j).B$. Denote the data matrix corresponding to each $\tilde{t}_k$ by T_k. If the sample is of size N, each T_k is an N-by-3 matrix. This means that for given parameter values B, the typical (i,j) element of the 5-by-5 covariance matrix of the 5-by-1 random vector $\tilde{T}.B$ can be computed as $B^T.T_i^T.T_j.B$, $\forall\, i,j$. The 3-by-3 matrix product inside the last expression, $T_i^T.T_j$, depends only on the data, for every $i,j = 1,2,3,4,5$. So while there are 25 such 3-by-3 matrix products, it is sufficient to compute each only once. Assumed primitive parameter values affect only the coefficient matrix B containing the shallow parameters β and λ. So even as we evaluate the objective function at 7.96 million sets of primitive parameter values, the 25 $T_i^T.T_j$ matrices stay the same.

This computational strategy is used in the R code for computing this variance-covariance matrix, which is then inverted to yield the optimal weight matrix. As a practical matter, we define 15 products of random variables, by multiplying each element of $\tilde{z}$ and $\tilde{x}^T$, to get $\widetilde{M} = vec(\tilde{T}^T)$. Let the corresponding data matrix be the N-by-15 matrix H. The 15-by-15 matrix of sums of squares and cross-products, $H^T.H$, can be partitioned into the 25 3-by-3 covariance matrices $T_i^T.T_j$.

Note that alternatively, we could, for each moment restriction k, $k = 1,2,3,4,5$, compute the realized value of the moment restriction for each data point and each set of primitive parameter values, as $T_{k,t}.B$, where

$T_{k,t}$ is the t^{th} observation in the data matrix T_k, and then compute the covariance matrix of the 5 moment conditions. While this is simpler to describe, it involves computing the covariance matrix for each set of primitive parameter values and is much more computer-intensive. By exploiting the structure of the moment conditions, we make do with a single computation of the covariance matrix $T_i^T . T_j$. This makes a very large task easily feasible. We still have one covariance matrix inversion for each set of parameter values.

<u>For the ancillary specifications (results tabulated in columns 2 and 3 of Table 9)</u>

To define the weight matrix for the ancillary specifications, we use the two different pricing errors corresponding to Regimes 3 and 2, and the associated available instruments as listed in Appendix D. This yields 9 moment conditions in all, 5 associated with Regime 3, and 4 with Regime 2. The corresponding 9-by-9 covariance matrix of these 9 moment conditions can conveniently be thought of as 2-by-2 = 4 blocks. Now we need to distinguish between variables from Regime 3 (with the superscript "(FT)") and Regime 2 (with the superscript "(T)"). A "T" without the parentheses denotes the transpose. For Regime 3, following exactly the discussion in the preceding subsection for the baseline specification, we can write the 5 moment conditions concisely (but restoring the superscript "(FT)")) as $E(\tilde{T}^{(FT)}. B^{(FT)}) = 0$, and the computed 5-by-5 covariance of these moment conditions will have a typical (i, j) element given by

$$(B^{(FT)})^T . (T_i^{(FT)})^T . (T_j^{(FT)}) . (B^{(FT)}), \forall\, i, j = 1, 2, 3, 4, 5$$

This is the $(1,1)$ block of the 9-by-9 covariance matrix. To get the 4-by-4 covariance matrix from just the Regime 2 moment conditions, which will yield the $(2,2)$ block of the 9-by-9 covariance matrix, note that analogous to the above discussion for the primary GMM specification based only on Regime 3, for Regime 2, we can define (using the superscript "(T)" to denote that only trading signals are available)

i) the instrument vector $(\tilde{z}^{(T)})^T = [\tilde{z}_1^{(T)}, \tilde{z}_2^{(T)}, \tilde{z}_3^{(T)}, \tilde{z}_4^{(T)}]$, where $\tilde{z}_1^{(T)} = 1$, $\tilde{z}_2^{(T)} = \tilde{\omega}^{(T)}$,
ii) random vector listing variables that help define the pricing error $(\tilde{x}^{(T)})^T = [\tilde{p}^{(T)}, \tilde{\omega}^{(T)}]$,
iii) the 4-by-2 random matrix $(\tilde{T}^{(T)}) \equiv (\tilde{z}^{(T)}). (\tilde{x}^{(T)})^T$ and
iv) parameters $(B^{(T)})^T = [1, -\lambda]$.

We can then define the 4 additional moment conditions concisely as $E((\tilde{T}^{(T)}). (B^{(T)})) = 0$. Define each 2-element row vector of the 4-by-2 matrix $\tilde{T}^{(T)}$ as $\tilde{t}_k^{(T)}, k = 1, 2, 3, 4$. Then the typical (i, j) element of the 4-by-4 covariance matrix of $(\tilde{T}^{(T)}). (B^{(T)})$ is given by $(B^{(T)})^T . Cov(\tilde{t}_i^{(T)}, \tilde{t}_j^{(T)}). (B^{(T)})$. Denote the data matrix corresponding to each $\tilde{t}_k^{(T)}$ by $T_k^{(T)}$. If the sample is of size N, each $T_k^{(T)}$ is an N-by-2 matrix. This means that for given parameter values $(B^{(T)})$, the typical (i, j) element of the 4-by-4 covariance matrix of the 4-by-1 random vector$(\tilde{T}^{(T)}). (B^{(T)})$ can be computed as

$$(B^{(T)})^T . (T_i^{(T)})^T . T_j^{(T)} . (B^{(T)}), \forall\, i, j.$$

The 2-by-2 matrix product inside the last expression, $T_i^T . T_j$, depends only on the data, for every $i, j = 1, 2, 3, 4$, and remains the same for any set of primitive parameter values, which will affect only the shallow parameter values.

This leaves only the two off-diagonal blocks, the 5-by-4 matrix that is the (1,2) block and its transpose the 4-by-5 matrix that is the (2,1) block. To compute the 5-by-4 matrix that is the (1,2) block, note that the typical (i,j) element is given by $(B^{(FT)})^T . Cov(\tilde{t}_i^{(FT)}, \tilde{t}_j^{(T)}) . (B^{(T)})$, and this would be computed using each 3-by-2 matrix $(T_i^{(FT)})^T . T_j^{(T)}$ from

$$(B^{(FT)})^T . (T_i^{(FT)})^T . T_j^{(T)} . (B^{(T)}), \forall\, i = 1, 2, 3, 4, 5, \text{ and } \forall\, j = 1, 2, 3, 4.$$

We showed above that for the 5-by-5 covariance matrix of the 5 moment conditions in the primary specification we could define 15 products of two random variables by multiplying each element of $\tilde{z}^{(FT)}$ and $(\tilde{x}^{(FT)})^T$, and then, given an N-observation data matrix H, partition the matrix with sums of squares and cross-products, $H^T . H$, into 25 blocks of 3-by-3 matrices. Similarly, for the 9-by-9 covariance matrix of the 9 moment conditions in the ancillary specification, we add 8 more products of two random variables (to the 15 products above) by multiplying each element of $\tilde{z}^{(T)}$ and $(\tilde{x}^{(T)})^T$.

We can then get a new random vector $\widetilde{M}$ (with 15+8=23 products of two random variables) by stacking $vec((\tilde{T}^{(FT)})^T)$ and $vec((\tilde{T}^{(T)})^T)$. Let the corresponding data matrix be the N-by-23 matrix H. The 23-by-23 matrix of sums of squares and cross-products, $H^T . H$, can be partitioned into 2-by-2 = 4 block matrices. The 25 3-by-3 covariance matrices $(T_i^{(FT)})^T . (T_j^{(FT)})$ will be defined using the 15-by-15 block matrix in the (1,1) position. The 16 2-by-2 covariance matrices $(T_i^{(T)})^T . (T_j^{(T)})$ will be defined using the 8-by-8 block matrix in the (2,2) position. To compute the 5-by-4 = 20 3-by-2 matrices $(T_i^{(FT)})^T . T_j^{(T)}$ we partition the 15-by-8 block in the (1,2) position. And transpose that to get the 4-by-5 = 20 matrices $(T_i^{(T)})^T . T_j^{(FT)}$, each of dimension 2-by-3, to fill in the (2,1) block.

Again, note that this covariance matrix has to be computed only once, even as we evaluate the GMM objective function at the many sets of primitive parameter values, each of which will only change the shallow parameters.

If we use the identity matrix as the weight matrix, then there is no covariance matrix to be computed or inverted to derive the weight matrix. And the initial evaluation of the GMM objective function at even 7.96 million points is very quick. It took less than half a minute. But when we used the resulting best starting points, they did not yield GMM estimates later with the optimal weight matrix that were as good as we got by using starting points obtained from evaluating the GMM objective function, by imposing the optimal weight matrix even at the time of each initial evaluation.