# Private Information and the Stock Market's Reaction to Earnings Announcements*

Clara Vega**

**Abstract**

This paper presents a new explanation for the post-announcement drift, i.e., the time-lag in market price adjustment to news surprises. Our explanation is that the post-announcement drift is a function of the private information agents have prior to news announcements. The more information people have about the "true" value of an asset – and the more they trade on this information – the smaller the abnormal return drift will be. We test this hypothesis empirically. Since private information is not observable by the econometrician, we use transaction-level data and a sequential trade microstructure model to calculate the probability of private information-base trading (PIBT) prior to an earnings announcement. We are able to show that a higher PIBT yields a lower post-announcement drift: the stocks with a low PIBT experience a drift that is 71% higher than high PIBT stocks.

Key Words: Earnings Announcements; Market Microstructure; High-Frequency Data; Expectations Data; Efficient Market Hypothesis; Private Information.
JEL: C1, D4, D8, G0

* I am grateful to the Bank of Spain for financial support. For useful comments I thank Luis Araujo, Celso Brunetti, Sean Campbell, Frank Diebold, Simon Gervais, Yuichi Kitamura, Frank Schorfheide, Dan Silverman and Vesna Straser, as well as participants in the October 2001 University of Pennsylvania econometrics seminar.

**Graduate Group in Economics, University of Pennsylvania, cvega@ssc.upenn.edu

## 1. Introduction

A key concern of financial economics has been the gap between theoretical efficient markets and what the data show about actual market performance. In an efficient market, security prices at any given time fully reflect all available information. A priori, there is good reason to believe that stock markets are efficient because such markets are paradigmatic examples of competition. Nevertheless, rather than adjusting immediately to news surprises, stock prices tend to drift over time in the same direction as the initial surprise. Over the last decade, researchers have provided three explanations for this phenomenon. First, that high transaction costs discourage traders from acting quickly on the information that has been announced. Second, that announcements with unexpected high (low) earnings make investing in firms more (less) risky, so the drift can be explained as a risk premium. Third, that agents fail to fully assess the implications of the information they have, and instead wait to react until there is a realization of the information through future earnings, or until analysts provide forecast revisions.

The first explanation shows why small firms, with corresponding high transactions costs, show higher drift than large firms. However it fails to explain why this drift is not eliminated by traders who face small transaction costs.[1] The second explanation seems insufficient because while there is a positive correlation between risk and earnings news, the difference in risk is not large enough to account for the post-announcement drift.[2] Finally, the third explanation would only account for the post-announcement drift observable in small stocks; as large firms tend to have an insignificant drift, suggesting traders do fully assess the implications of earnings information and act accordingly.

We propose a fourth explanation, which reconciles the empirical evidence for small as well as large firms. We conjecture that the announcement drift is a function of the private information (PI) agents have prior to the official earnings announcement. So, independent of transaction costs and firm size, the more information people have about the "true" value of an asset –and the more they trade on this information– the smaller the abnormal return drift the asset

---

[1] Some traders face no broker commissions and can bypass the specialist's bid-ask spread, which are the two major transactions costs.

[2] Bernard and Thomas (1989), Fama and French (1996).

value will experience. We test this hypothesis empirically. Since PI is not observable by the econometrician, we use transaction-level data and a sequential trade microstructure model to calculate the probability of private information-base trading (PIBT) prior to an earnings announcement. We are able to show that a higher PIBT yields a lower post-announcement drift: the stocks with a low PIBT experience a drift that is 71% higher than high PIBT stocks.

Researchers in economics and finance have long studied the effect of public announcements and private information on prices, volume and volatility.[3] In a recent paper Hong and Stein (1999) introduced boundedly rational agents, called "newswatchers," who react to the same news identically. However, in this study, each newswatcher receives a piece of PI about a firm but does not know the information held by others. Assuming that each newswatcher's valuation depends on their own portion of PI, and not on observed prices, they generate an equilibrium wherein prices under-react to initial news and gradually drift towards the initial reaction. The drift emerges as agents aggregate PI about the firm over time. As noted by Hong and Stein, the acquisition of different signals by different agents, in response to public announcements, is rather unlikely. They suggest that the drift following public news announcements might be caused by a lack of consensus among agents. Instead of assuming that prices adjust slowly as individuals aggregate different pieces of PI received over time, Hong and Stein suggest that individual investors have private interpretations of the information revealed in news announcements.

Garfinkel and Sokobin (2000) and Kandel and Pearson (1995) provide empirical evidence supporting Hong and Stein's theory. Both studies use volume as a proxy for consensus among informed traders and show that higher than expected trading volume around an earnings announcement is associated with more drift. The sequential trade model we use does not include a consensus variable. Agents in this model sometimes receive private information, and when they do, there is unanimous consensus among informed traders. Nevertheless, we can associate a

[3] Empirically oriented papers include Bernard and Thomas (1989), Joel Hasbrouck (1991), Maddala and Nimalendran (1995), Kandel and Pearson (1995), Fama and French (1996), Fama (1970, 1998) and Garfinkel and Sokobin (2000). On the theoretical side, see Copeland and Galai (1983), Kyle (1985), Admati and Pfleiderer (1988) , Barberis, Shleifer and Vishny (1998), and Hong and Stein (1999).

high probability of information-based trading prior to the announcement with a higher consensus among traders. Indeed, our estimate of private information-based trading is negatively correlated with volume. An interesting result is that both variables, volume and the probability of informed trading, are important when predicting the size of the post-announcement drift.

The study proceeds as follows. Section 2 describes the sequential trading microstructure model we use; it allows us to calculate the probability of PIBT taking place, and to predict what happens after an earnings announcement. Section 3 describes the three data sets being used, in particular the transaction-level data and how we use it to extract the number of daily buys and sells. In this segment, we also present a cross-section as well as a time-series of the estimated parameters. Section 4 relates the information trading variable to post-announcement abnormal returns. We conclude in section 5.

**2. Private Information**

Private information trading is unobserved by the econometrician. Our approach is to use a proxy for this unobserved variable. The proxy we have in mind is Easley, Kiefer, O'Hara and Paperman's (EKOP) (1996) private information estimate.[4] Using this proxy serves two purposes: first, it allows us to estimate how much private information agents have prior to an earnings announcement and secondly it provides the economic theory we need to understand why the post-announcement drift can be written as a function of private information. What follows is a brief description of their model.

There are three players in this game, liquidity traders, informed traders and a market maker. All players are risk neutral, they have infinite wealth, there are no transactions cost and there is no discounting by traders. The implications of the model do not depend on the agents' risk neutrality, yet risk neutrality simplifies the problem. This assumption, in addition to a competitive market assumption, forces the market maker to set prices equal to his expected value of the asset, conditional on his information set. The infinite wealth assumption allows us to disregard budget constraints. The no-transactions cost assumption is a sufficient condition for market efficiency, and the no-discounting assumption is reasonable since agents are optimizing

---

[4]Their model belongs to the Glosten and Milgrom (1985) sequential trade class of models.

their behavior over one day.

Liquidity traders trade for exogenous reasons to the model and they arrive to the market, according to an independent Poisson distribution, with a daily arrival rate equal to $\varepsilon_b$ for liquidity buyers and $\varepsilon_s$ for liquidity sellers.[5] Informed traders trade for speculative reasons. They receive information, according to an independent Poisson distribution, with arrival rate equal to $\mu$; hence, their arrival rate to the market is equal to $\mu$ as well. Prior to the beginning of each trading day *d*, nature decides whether informed traders receive private information or not. On average, they receive private information with probability $\alpha$. These information event is either good news with probability $\delta$ or bad news with probability $1-\delta$. In equilibrium, informed traders buy one unit of the asset if they receive a good news signal or sell one unit of the asset if they receive a bad news signal.

The market maker sets prices to buy or sell at each time *t* in [0,*T*] during the day, and then executes orders as they arrive. He observes the trade (either a buy or a sell), and uses this information to update his beliefs. New prices are set, trades evolve, and the price process moves in response to the market maker's changing beliefs. This process is captured in figure 1, taken from EKOP. At the first node of the tree, nature selects whether an information event occurs. If an event occurs, nature then determines if it is good news or bad news. The three nodes (no event, good news, and bad news) before the solid vertical line in figure 1, occur only once per day. Then, given the nodes selected for the day, traders arrive according to the relevant Poisson process. That is, on good event days, the arrival rates are $\varepsilon_b+\mu$ for buy orders and $\varepsilon_s$ for sell orders. On bad event days, the arrival rates are $\varepsilon_b$ for buy orders and $\varepsilon_s+\mu$ for sell orders. Finally, on non-event days, only uninformed traders arrive and the arrival rate of buys is $\varepsilon_b$ and of sells is $\varepsilon_s$.

Over time the process of trading, and learning from trading, results in prices converging to full information levels. The higher PI is the faster this convergence will be. We can view informed traders as traders who know the "true" value of the asset. They can either be insiders or traders who are particularly skillful in processing public information. This is why PI in our

[5]We estimated the model assuming that the arrival rate of liquidity buyers and sellers is the same, but a log likelihood ratio test indicates that different arrival rates fit the data better.

model is not "bad," on the contrary it helps the market to be efficient. If informed traders have not received private information before the public announcement, then, we are more likely to see an adjustment in stock prices after the public news announcement.

To illustrate how the theoretical model can predict what happens after an earnings announcement we simulate the model assuming public announcements are a noisy signal of the true value of the asset. We refer you to the appendix to see the results of the simulation. In the next sub-section we explain how we calculate PI using real data.

Likelihood Function

As shown by EKOP, PIBT is estimated by dividing the estimated arrival rate of informed trades by the estimated arrival rate of all trades,

$$\hat{PI} = \frac{\hat{\alpha}\hat{\mu}}{\hat{\alpha}\hat{\mu} + \hat{\varepsilon}_b + \hat{\varepsilon}_s}.$$

This is the variable we use to proxy for the level of information gathering in our tests. In order to estimate this variable, we first need to estimate the parameters $\theta = \{\varepsilon_b, \varepsilon_s, \mu, \delta, \alpha\}$. To do so we maximize the following likelihood function,

$$L(S,B|\theta,\psi) = \alpha\delta L(S,B|\theta,\psi=-1) + \alpha(1-\delta) L(S,B|\theta,\psi=1) + (1-\alpha) L(S,B|\theta,\psi=0)$$

this is the unconditional probability that B number of buys and S number of sells occur at a given time, which is a mixture of three conditional probabilities, the conditional probability that we observe S sells and B buys given a bad news signal, a good news signal or a no news signal. The conditional probability that we observe S sells and B buys, given a bad news signal, is

$$L(S,B|\theta,\psi=-1) = e^{-\varepsilon_b T}\frac{(\varepsilon_b T)^B}{B!} e^{-(\varepsilon_s+\mu)T}\frac{((\varepsilon_s+\mu)T)^S}{S!},$$

which is a Poisson Process with the number of buys arriving at a rate equal to $\varepsilon_b$, and the number of sells arriving at a rate equal to $\varepsilon_s + \mu$. The conditional probability that we observe S sells and B buys, given a good news signal, is

$$L(S,B|\theta,\psi=1) = e^{-(\varepsilon_b+\mu)T}\frac{((\varepsilon_b+\mu)T)^B}{B!} e^{-\varepsilon_s T}\frac{(\varepsilon_s T)^S}{S!}.$$

Finally, the conditional probability that we observe S sells and B buys, given a no news signal, is

$$L(S,B|\theta,\psi=0) = e^{-\varepsilon_b T}\frac{(\varepsilon_b T)^B}{B!} e^{-\varepsilon_s T}\frac{(\varepsilon_s T)^S}{S!}.$$

We set *T = 1*, so that arrival rates are interpreted as *daily* arrival rates. Assuming days are independent, the likelihood of observing the data $M = \{B_d, S_d\}_{d=1}^{D}$ over D days is just the product of the daily likelihoods,

$$L(M|\theta,\psi) = \prod_{d=1}^{D} L(\theta,\psi|B_d,S_d) \qquad (1)$$

The likelihood function is differentiable, but there is no closed form solution to the five ($\varepsilon_b, \varepsilon_s, \mu, \delta$ and $\alpha$) first order conditions. Nevertheless, we can look at a simulation to see what the likelihood function is doing. We generate three different samples of forty days, where buys and sells are generated as normally, independently and identically distributed variables with mean equal to forty and standard deviation equal to three. In the first sample labeled "good news" we choose at random sixteen days out of the forty days and we add twenty buys more to the "usual" level of buys. The "bad news" sample consists of sixteen days that have twenty sells more than "usual". Finally, in the sample "both news" we choose at random sixteen days out of the forty days and we add twenty buys more to the "usual" level of buys, simultaneously and independently we choose at random another set of sixteen days and we add twenty sells more to the "usual" level of sells. Next, we calculate the likelihood function in each case. The parameter estimates are shown in table 1. The likelihood function sets $\varepsilon_s$ and $\varepsilon_b$ equal to the mean sells and buys of the underlying generating process. The parameter $\mu$ is set equal to twenty, the "abnormal" number of buys or sells. The parameter $\alpha$ is equal to the proportion of days where there were an "abnormal" level of transactions, in our case sixteen days out of forty days or forty percent of the time. The parameter $\delta$ is set equal to zero on the pure "good news" sample, equal to one on the pure "bad news" sample and equal to fifty percent on the "both news" sample.

**3. Data and Parameter Estimates**

To maximize the likelihood function described above we need to know the number of buys and sells that occur each day. We use the algorithm developed by Lee and Ready (1993) to calculate the daily number of buys and sells.[6] This algorithm compares transaction prices to the

---

[6]Ellis, Michaely and O'Hara (1999) evaluate how well the Lee and Ready algorithm performs and they find that the algorithm is 81.05% accurate. Optimally we would like to have the actual trade direction. Unfortunately this data is only available for very short sample periods, for one or two months. So if we want to use a long sample period, we are forced to use such algorithm. Nevertheless this

mid-quote five seconds before the transaction takes place. If the transaction price is above the mid-quote then we classify the trade as a buy, if the transaction price is below the mid-quote we classify the trade as a sell. If the transaction price is equal to the mid-quote then we use the "tick test," that is if the transactions price is higher than the previous price we classify the trade as a buy and if it is below the previous transactions price we classify it as a sell. We use bid quotes, ask quotes and transactions prices from the Trade and Quotes (TAQ) database.

To calculate the abnormal (size-adjusted) returns we use the Center for Research in Security Prices (CRSP) data,

$$AR_{kd} = R_{kd} - R_{pd},$$

where $AR_{kd}$ is the abnormal return for firm *k*, day *d*. $R_{kd}$ is the raw return for firm *k* day *d* and $R_{pd}$ is the portfolio return for portfolio *p* for day *d*. The portfolio includes all NYSE firms that belong to the same size decile firm *k* is a member of at the beginning of the calendar year.[7] Firm size is measured by the market value of common equity. To calculate the standardized surprise we use actual earnings announcement data (primary earnings per share, excluding extra-ordinary items, adjusted for stock splits and stock dividends) and a forecast of the earnings announcement data. Both of these figures are taken from I/B/E/S international Inc. We calculate the surprise for stock *k* on day *d* as follows,

$$S_{kd} = \frac{A_{kd} - E_{kd}}{\hat{\sigma}_{kd}},$$

where $A_{kd}$ is the announced stock *k* earnings, $E_{kd}$ is the market expected value as distilled in the I/B/E/S median forecast, and $\hat{\sigma}_{kd}$ is a rolling estimate of the standard deviation of prior forecast errors where the rolling window is four years. We use past forecast errors to compute the standard deviation, as opposed to the entire sample, to avoid hindsight biases.

Our sample period starts on January 1, 1993 and ends on December 31, 1999. We start on January 1, 1993 because this is the first year TAQ data is available. In order to compute $\hat{\sigma}_{kd}$ we use I/B/E/S data going back to January 1, 1989. We choose all stocks that are traded in the

---

algorithm is the one that performs best.

[7] We compute the abnormal returns in this way to control for the Banz-Reinganum size effect, namely low (high) market capital firms have higher (lower) mean return than the market portfolio.

NYSE, because the microstructure model we use is compatible with this trading environment.[8] Out of this set of stocks, we only include those stocks whose earnings announcements are forecasted by I/B/E/S. We exclude stocks that did not have at least four years of earnings announcement data prior to including the stock in the sample (because we need four years of data to compute $\hat{\sigma}_{kd}$). We exclude stocks in any year which did not have at least 40 days with quotes or trades, prior to the earnings announcement. Finally, merging the three data sets together we are left with a sample of stocks ranging from 746 to 540 to be analyzed each quarter.

Post-Announcement Drift Stylized Facts

Ball and Brown (1968) discovered the positive correlation between earnings announcement surprises and cumulative abnormal returns over sixty days following the earnings announcement. Many people since then have studied this relationship. In particular, Bernard and Thomas (1989) studied this relationship in the sample period between 1974 to 1985. In our sample period from 1993 to 1999, the pattern still holds. In table 2 we show the means of variables, categorized by surprise deciles. The surprise decile is a ranking from 1 to 10 of this quarter's surprise $S_{kd}$ (firm $k$, day $d$), based on the distribution of earnings surprises in the *previous* quarter. Again, we do this to prevent hindsight biases. The two-day abnormal return is the cumulative size adjusted return on the day of the announcement and the day after. This is a measure of the contemporaneous effect of the earnings announcement on the abnormal return. We look at the day after, because some earnings announcements take place after the stock market is closed and the price reaction is observed the next day. Drift is calculated as the sum of size adjusted returns over trading days [+2,+40], where day 0 is the day after the earnings announcement date.

The results in table 2 are consistent with past evidence. The two-day abnormal return is a monotonic function of the earnings surprise. The post-announcement drift is close to a monotonic function, with the exception of the third decile, which is associated with a bigger drift than the drift associated with the first and second deciles. In general it appears that our sample does not differ substantially from past samples in papers examining drift, like Bernard and

[8]The NYSE is a specialist market (it uses market makers) and it is a competitive market, because the market makers faces competition from the limit order book.

Thomas (1989).

In figure 2 we reproduce the result of Bernard and Thomas (1989) with a different sample period and a smaller set of stocks.[9] Figure 2 is a plot of the cumulative size adjusted returns over trading days [- 40, +1] and over trading days [+2, +40]. This figure is qualitatively and quantitatively similar to Bernard and Thomas figure 2. The only two differences are the aforementioned third decile post-announcement drift outlier and that their figure is more symmetric around a zero announcement surprise than ours. These two differences can be attributed to our under-sampling of stocks on the negative side of the earnings announcement surprise. Two factors contribute to our under-sampling. The first factor is the stock sample size. Bernard and Thomas sample size is on average 1,600 stocks per quarter. Our stock sample size is smaller; it varies from 746 to 540 stocks per quarter because we restrict our attention to stocks traded in the NYSE and stocks that appear in all three data sets (CRSP, TAQ and I/B/E/S). The second factor is the sample period. In their sample period there are expansionary periods as well as recessionary periods. Whereas our sample period is purely an expansionary period. Nevertheless, what we want to explain is the positive correlation between the announcement surprise and the post-announcement drift, which is present in our sample period. Moreover the difference in the forty day post-announcement drift between the first and tenth decile is 2.94%, which is economically significant. If we were to re-balance our portfolio every 40 days after an earnings announcement with a short position in the first decile stocks and a long position in the tenth decile stocks we would obtain an 18% annual return compared to the annualized average return of the SP500 index in our sample period equal to 16.8%.

[9]Their pre-announcement and post-announcement periods are over 60 trading days. We restrict our period to 40 trading days because even though on average there are 60 trading days between earnings announcements there are some instances where earnings announcements are separated by as little as 40 trading days. Since we do not want to overlap earnings announcement days we restrict the sample period to 40 days.

Parameter Estimates

We use forty days of buys and sells data before each earnings announcement, for all stocks, to maximize the likelihood function given in equation (1) with respect to the parameter space $\theta = \{\varepsilon_b, \varepsilon_s, \mu, \delta, \alpha\}$. We allow the parameters to vary across stocks and across time, so we estimate a separate likelihood function for each stock, each period.

Cross-Section Variation of PI

Since information trading is not observable, it is natural to ask whether this measure of private information is a good proxy. To explore this issue we look at the cross-section and time-series properties of PI. Exploring the properties of PI in both dimensions will also help us understand the explanatory power of PI in the cross-section regressions, time-series regressions and pooled data analysis. In figure 3 we show kernel estimates of the density of the coefficient of variation of PI across time and across stocks.[10] The coefficient of variation across time is calculated by fixing a stock and dividing the standard deviation of PI by the mean of PI across time. To calculate the coefficient of variation across stocks we fix a time period and divide the standard deviation of PI by the mean of PI across stocks. This plot, along with the statistics provided below, show that the variation of PI across stocks is larger than the variation across time. The difference between the two medians is statistically significant.[11] This is why we can think of PI as more or less a stock "characteristic," as Easley, Hvidkjaer and O'Hara (2000) suggest.

In particular, we can show that a high PI is characteristic of small firms. In table 4 we show the pooled data simple correlation matrix of PI against, among other variables, firm capital size. The firm capital size is a cardinal number from 1 to 10 corresponding to deciles based on January 1 market value of equity for all NYSE firms. The decile ranking changes on an annual bases. The correlation is significantly negative, -0.50, with a cross-sectional range over the 27 periods of -0.67 to -0.3 implying that across stocks within the same quarter PI is negatively

---

[10] We use the Epanechnikov kernel as a weighting function and we use Silverman's bandwidth selection criteria.

[11] We perform a Mann-Whitney test, under the null hypothesis of equal sample medians, we reject the null hypothesis at a 5% confidence level. For more details on this test we refer the reader to Sheskin (1997).

correlated with the firm capital size. To convince ourselves, we present in figure 4 the kernel estimates of the density of PI for "large" firms and "small" firms from 1993:01 to 1997:04.[12] We exclude periods 1998:01 to 1999:03 to conserve space, nevertheless in table 5 we show a complete set of statistics. At first sight, we can say that the probability distribution of PI for "large" firms has a significantly smaller median than the probability distribution of PI for "small" firms. To test this hypothesis formally we present in table 5 the Mann-Whitney tests statistics for all periods. We test the hypothesis that the two samples have the same sample median, against the alternative hypothesis that they have different medians. The critical value at the five percent confidence level is 1.96, since all the numbers shown in the table are above this number, we can soundly reject the null hypothesis in favor of the alternative for all the periods.[13]

The negative relationship between private information and firm size is consistent with Diamond and Verrecchia (1991) who assert that asymmetric information is largest for small firms. The logic behind their claim is that large firms have an incentive to disclose more information, whereas small firms do not benefit as much from information disclosure. The negative relationship between firm size and asymmetric information is widely accepted in the literature, so the fact that this relationship holds with our private information proxy is good news.

This strong negative correlation is of additional interest to us because Bernard and Thomas (1989) documented an inverse relationship between post-announcement drift and firm size. This negative relationship holds true for our sample period. As table 3 shows, the drift for large firms, although statistically significant is much smaller than the drift for small firms. The economic explanation for this relationship is transactions costs. Large stocks have lower transactions cost than small stocks. Since transactions costs prevent people from trading on the information they have, we expect the market for large stocks to be more efficient or to have a smaller post-announcement drift. Since our explanatory variable is negatively correlated with the firm size, some of the explanatory power of our variable comes from its negative correlation with

---

[12] "Large" capital firm size refers to firms in the tenth decile and "small" capital firm size refers to all the other stocks. This definition makes sense given our stock sample.

[13] In our sample period, 1996:04 jumps out as the period where PI for large firms is closest to PI for small firms. Even though we cannot accept the null that the two sample medians are the same it would be interesting to find out what was happening in the stock market.

firm size. To control for this effect we must include in our regressions firm size as one of the right hand side variables.

To further examine the validity of our proxy we look at the correlation between PI and other variables that are related to information-based trading like volume and bid-ask spreads. We calculate the bid-ask spread as the average daily spread over forty days prior to each earnings announcement, for each stock. We adjust the spread by the average price of the stock over the previous year. Volume is defined as the logarithm of the average daily dollar trading volume over forty days prior to each announcement adjusted by the average daily number of shares outstanding over the previous year. Trading volume is correlated with the number of shares outstanding, so we adjust the trading volume to control for cross-sectional differences. Both of these variables are calculated using CRSP data. According to previous literature the spread has three components, an adverse selection costs, an order-processing costs and an inventory holding costs. Holding the last two components constant, we expect PI to be positively correlated with spread, because a high PI means that the market maker incurs a higher adverse selection cost. Indeed, the correlation, shown in table 4, is significantly positive, 0.346, with a cross-sectional range over the 27 periods of 0.58 to 0.14 implying that across stocks within the same quarter PI is positively correlated with the bid-ask spread.

High trading volume is associated with a low consensus among informed traders.[14] Our model assumes that when traders receive private information, they all agree on its meaning, in other words there is perfect consensus among informed traders. So we expect PI and volume to be negatively correlated. The correlation, shown in table 4, is significantly negative, -0.587, with a cross-sectional range over the 27 periods of -0.72 to -0.37 implying that across stocks within the same quarter PI is negatively correlated with volume. So far our latent variable seems to be a good proxy for private information-based trading.

<u>Time Section Variation of PI</u>

PI is a function of the parameters $\varepsilon_b, \varepsilon_s, \mu, \text{and}\, \alpha$. In figure 5, we plot a time series of these parameter estimates plus our estimate of the probability of bad news, $\delta$, averaged across all stocks

[14]See, for example, Kandel and Pearson (1995) and Garfinkel and Sokobin (2000).

in our sample. The parameters $\varepsilon_b, \varepsilon_s$ and $\mu$ are not stationary. These parameters are related to the trading frequency, hence they are trending upwards as the number of transactions on the NYSE has increased over the years. On the other hand, the estimates of $\delta, \alpha$ and PI are stationary over the years. It is interesting to notice that the probability of a bad news event, $\delta$, is consistently below 0.5. The time period we study is an expansionary period, so it makes sense that most of the private information traders received in this period were good news events.

In figure 5 we cannot see whether particular stocks experience significant changes in PI over time. It is counterproductive to show a plot of all 746 stocks, but we can tell the story of a particular stock and see whether it fits the general case. The particular stock we pick is Exxon Corporation. We plot a time series of its parameter estimates in figure 6. Similar to figure 5, the parameters $\varepsilon_b, \varepsilon_s$ and $\mu$ are not stationary, but the estimates of $\delta, \alpha$ and PI are stationary over the years. On January 25 1994, Exxon experienced the largest, relative to its history of earnings, earnings announcement surprise in our sample period. Prior to this announcement the probability of information trading was in the top 18 percentile. PI was equal to 0.1, the second spike in PI shown in figure 6. If our model were true, we would predict low abnormal returns at the time of the earnings announcement and no post-announcement drift, because PI is very high, regardless of the earnings announcement surprise size. Indeed, the abnormal return at the time of the announcement was in the bottom 22 percentile and the post-announcement drift was in the bottom 18 percentile, as well, rather than in the top 18 percentile as the size of the earnings announcement surprise would have predicted. Of course, this is only one example, the next section will test our theoretical model using time-series and pooled data regressions.

One question of interest the previous example brings up is whether we need to estimate time varying parameters. The fact that $\varepsilon_b, \varepsilon_s$ and $\mu$ are not stationary does not mean we need to estimate time varying parameters. The parameters $\varepsilon_b, \varepsilon_s$ and $\mu$ are not stationary because the number of buys and sells are trending upwards. So to fix this problem we simply need to de-trended the number of buys and sells and then estimate the parameters. What we are interested in is whether our parameters are time varying once the trend has been taking care off. In other words, whether a particular stock parameters change significantly conditional on the environment. This subject is of interest to us, because we would like to explain the post-announcement drift with a variable that is varying across time, as well as across stocks, since the post-announcement drift changes across time as well as across stocks. To test whether time-varying parameters are important or not, we perform a likelihood ratio test.

We calculate the constrained maximum likelihood estimator, $\hat{\theta}_c = \{\hat{\varepsilon}_b, \hat{\varepsilon}_s, \hat{\mu}, \hat{\delta}, \hat{\alpha}\}$, for each stock by maximizing equation (1) constraining the parameters to be the same across time. We also calculate the unconstrained maximum likelihood estimator, $\hat{\theta}_u$, for each stock by maximizing equation (1) allowing the parameters to vary over time. To test the hypothesis we further restrict ourselves to the sample period from 1993 to 1994 where we reject the null hypothesis of parameters $\varepsilon_b, \varepsilon_s, \text{and}\, \mu$ trending for 70% of the stocks. Naturally, the likelihood function evaluated at its optimal unconstrained value, $\hat{\theta}_u$, should be bigger than or equal to the likelihood function evaluated at is optimal constrained value, $\hat{\theta}_c$. Nevertheless, we can say that the constraint model fits the data well if the difference between the two likelihoods is not very big. Formally, we test the null hypothesis $\theta = \hat{\theta}_c$ using a likelihood ratio test. The likelihood ratio test statistic is equal to $-2(\ln\hat{L}_c - \ln\hat{L}_u)$, where $\hat{L}_c$ is the likelihood function described in equation (1) evaluated at $\hat{\theta}_c$ and $\hat{L}_u$ is the same likelihood function evaluated at $\hat{\theta}_u$. The large sample distribution of this test statistic is a chi-square distribution with the number of degrees of freedom equal to the number of restrictions. Since we have too many stocks it is not very practical to present the likelihood ratio test for all stocks. Instead we present in table 6 the fraction of stocks for which we accept the null hypothesis when we constrain our parameters over two quarters, three quarters, four quarters, five quarters and six quarters.

Table 6 shows the fraction of stocks whose parameters do not vary significantly over the sample period. Even if we only constrain the parameters over a two quarter period the unconstrained model fits 70% of the stocks better, this percentage is larger than the 30% of stocks whose parameters are trending upwards. So it is safe to say that the time varying parameters are important, once we get rid of the trend. As we expected the percentage of stocks the constrained model fits well is decreasing as we increase the constrained length. By the time we reach a length of six quarters, less than 1 percent of the stocks have time invariant parameters.

## 4. Empirical Results

We want to test the following two hypothesis: Does the probability of PIBT prior to earnings announcements decrease or eliminate the two-day abnormal return? Does the probability of PIBT prior to earnings announcements decrease or eliminate the post-announcement drift? To test these two hypothesis we model the size of the two-day abnormal returns, $|AR_{kd}|$, as a linear function of the probability of PIBT in stock *k* prior to the announcement, $PI_{kd}$,

$$|AR_{kd}| = \alpha_p + \beta_p PI_{kd} + \varepsilon_{kd} \tag{2}$$

and we test the null hypothesis of whether the coefficient in front of PI is negative, $H_0\!:\beta_p<0$. We also model the size of the post-announcement cumulative abnormal returns over forty days, trading days [+2, +40], $|CAR_{kd}|$, as a linear function of $PI_{kd}$,

$$|CAR_{kd}| = \alpha_p + \beta_p PI_{kd} + \varepsilon_{kd} \tag{3}$$

and we test the null hypothesis of whether the coefficient in front of PI is negative, $H_0\!:\beta_p<0$. First, we look at a time-series analysis. This analysis will tell us whether a given stock's PI significantly varies across time and whether this variation can explain a particular stock's time-varying post-announcement drift. Then we look at a cross-section analysis. This analysis will tell us whether PI variations across stocks can explain the post-announcement drift differences across stocks. Finally, we look at a pooled data analysis to see how both, time-variation and cross-sectional variation in PI can explain the time-varying and cross-varying post-announcement drift.

Time-Series Analysis

Equations (2) and (3) in a time-series analysis become,

$$|AR_d| = \alpha_p + \beta_p PI_d + \varepsilon_d \tag{4}$$

$$|CAR_d| = \alpha_p + \beta_p PI_d + \varepsilon_d \tag{5}$$

We run these regressions separately for 656 stocks that have at least 20 observations in our sample period. A priori, we do not expect very good results, because Maddala and Nimalendran (1995), who measure the effect earnings surprises have on the two-day abnormal return, do not find the time-series results very promising and move on to a pooled data analysis.[15] Nevertheless, we can learn something from the time-series results presented in table 7. Before we talk about the results, note that we use a bootstrap method to derive the variance-covariance estimator, with which we compute the t-statistics. We use the bootstrap method to control for the fact that we use a two-step procedure: first we compute $PI_{kd}$ using the likelihood function given in equation (1) and second we estimate equation (4). The bootstrap procedure also takes care of the small sample bias induced because we use 40 observations to compute $PI_{kd}$ each time. We use this method throughout our empirical analysis.

Table 7 shows the fraction of stocks whose $\beta_p$ coefficient in equation (4) is statistically significant at a 5% significance level. This fraction is 13.10 %, like we expected from Maddala and Nimalendran's empirical results this is not a very high fraction. Nevertheless, 60.70 % of the coefficients were negative and 12.48 % were statistically significant and had the expected sign, i.e. $\beta_p<0$. This means that most of the statistically significant coefficients had the expected sign. We can compare these percentages with the ones we obtain when we run equation (4) using the earnings announcement surprise as the right hand side variable like in Maddala and Nimalendran's study,

$$|AR_d| = \alpha_s + \beta_s |S_d| + \varepsilon_d \tag{6}$$

$$|CAR_d| = \alpha_s + \beta_s |S_d| + \varepsilon_d \tag{7}$$

The fraction of stocks whose $\beta_s$ coefficient in equation (6) is statistically significant is 14.85 %, a bit higher than the percentage of $\beta_p$ coefficients that are statistically significant. The percentage of positive $\beta_s$ coefficients is 64.63 %. A positive coefficient is the expected sign according to Bernard and Thomas (1989) paper and the stylized facts shown in table 2. The percentage of statistically significant coefficients with the expected sign, i.e. $\beta_s>0$ is 13.97 % . When we put both variables as right hand side variables,

[15]Maddala and Nimalendran treat the earnings announcement surprise as a latent variable. They use the unobserved components approach that treats earnings surprises as an unobserved variable that occurs as an explanatory variable in several equations. They also use earnings announcement surprises (using I/B/E/S forecasts as the expected value of the earnings) and price changes as proxies for the unexpected component of the earnings

$$|AR_d| = \alpha + \beta_p PI_d + \beta_s |S_d| + \varepsilon_d \quad (8)$$

$$|CAR_d| = \alpha + \beta_p PI_d + \beta_s |S_d| + \varepsilon_d \quad (9)$$

The percentages do not change much. This means that, in the time-series analysis, the earnings announcement surprise is a better predictor of the two-day abnormal return, than our PI variable, nevertheless our PI variable does relatively well.

These results change when we look at the post-announcement drift. In the time-series analysis, PI is a better predictor of the post-announcement drift, than the earnings announcement surprise. The fraction of stocks whose $\beta_p$ coefficient in equation (5) is statistically significant is 13.10 % as compared to 11.35%, the fraction of stocks whose $\beta_s$ coefficient in equation (7) is statistically significant. We compare 72.93 % of the $\beta_p$ coefficients in equation (5) were negative to 56.29 % of the $\beta_p$ coefficients in equation (7) were positive. Finally, 12.23% of the $\beta_p$ coefficients in equation (5) were negative and statistically significant compared to 8.24% of the $\beta_s$ coefficients in equation (7) were positive and statistically significant.

Cross-Section Analysis

Equations (2) and (3) in a cross-section analysis become,

$$|AR_k| = \alpha_p + \beta_p PI_k + \varepsilon_k \quad (10)$$

$$|CAR_k| = \alpha_p + \beta_p PI_k + \varepsilon_k \quad (11)$$

We run these regressions separately for each period from the first quarter of 1993 through the third quarter of 1999. In total we run 27 regressions each of which has at least 540 observations and no more than 746. We present these results in table 8. As Maddala and Nimalendran point out the cross-sectional error variances of regression models similar to equation (10), where the right hand side variable are volume, spread and earnings surprises rather than PI, are proportional to the inverse of the square of the market value. In other words we have heteroskedasticity. This is case for both equations (10) and (11), where we find the absolute value of both errors in regressions (10) and (11) to be significantly related to the inverse of the market value. To fix the heteroskedasticity and obtain efficient standard error estimates we weight our variables by the market value capital decile, $C_{kd}$ (a cardinal number from 1 to 10 corresponding to small, medium and large firms). In other words we run a weighted least squares regression. The heteroskedasticity is present in all regressions that follow. So we run weighted least square regressions from now on.

The results shown in table 8 are qualitatively similar to the time-series results but they are very different quantitatively speaking. The percentages are 2 to 5 times larger in the cross-sectional analysis. Nevertheless, the increase in order of magnitude is due to the fact that we have increased the number of observations we use in the regression from 27 to 746. The conclusion is similar to that in the time-series analysis. The earnings announcement surprise is a better predictor of the two-day abnormal return, than our PI variable. Although the PI variable does better in the cross-sectional analysis compared to the earnings surprise variable than in the time-series analysis. In the cross section we also find that PI is a better predictor of the post-announcement drift, than the earnings announcement surprise.

Going back to the two hypotheses we wanted to test. In the time-series and cross section analysis $\beta_p$ is negative more than 50% of the time, but we will formally test the null hypothesis, $H_0: \beta_p < 0$, in the pooled data analysis, where we include other variables that are able to explain the post-announcement drift and the two-day abnormal return.

Pooled Data Analysis

The pooled data analysis results are presented on table 9 through table 15. It is worth noting that the coefficient estimate, $\beta_p$, is bias towards zero due to the errors in variables problem. So the results we present have a potentially larger coefficient than the one we present. The equations we have in mind are,

$$|AR_{kd}| = \alpha + \beta_p PI_{kd} + \beta_s |S_{kd}| + \varepsilon_{kd} \tag{12}$$

$$|CAR_{kd}| = \alpha + \beta_p PI_{kd} + \beta_s |S_{kd}| + \varepsilon_{kd} \tag{13}$$

In table 9 we present the coefficient estimate of equation (12) and (13) including one independent variable at a time and both independent variables at the same time.

First, we focus in the case where we only include PI as an independent variable. In this case, we accept the null hypothesis $H_0: \beta_p < 0$. As a robustness check we divide the cumulative abnormal returns over the first 20 days [+1, +20], $|CAR_{1kd}|$, and the last 20 days [+21, +40], $|CAR_{2kd}|$,

$$|CAR_{1kd}| = \alpha + \beta_p PI_{kd} + \beta_s |S_{kd}| + \varepsilon_{kd} \tag{14}$$

$$|CAR_{2kd}| = \alpha + \beta_p PI_{kd} + \beta_s |S_{kd}| + \varepsilon_{kd} \tag{15}$$

The negative coefficient confirms that our theory is able to predict a lower drift when PIBT is high. Since the magnitude of the $\beta_p$ coefficient in equation (15) is larger than in equation (14), we can say that PI has a stronger effect on the second half of the drift than on the first half, although the coefficient difference between equation (14) and (15) is not statistically significant. This result is very interesting, because figure 1 shows a slight reversal of the drift in the second half of post-announcement forty-day sample.

Now we focus in the case where we only include the absolute value of the earnings surprise as the independent variable. As we expected from the time-series and cross-section analysis the $\beta_s$ coefficient in table 9 is positive and statistically significant. This also confirms the results presented in table 2. It is interesting to notice that the $\beta_s$ coefficient loses its significance in the second half of the drift. Like we mentioned before, figure 1 shows a slight reversal of the drift in the second half of the post-announcement forty-day sample and we do not expect the size of the earnings surprise to be able to predict this reversal.

Finally, we include both variables, PI and the absolute value of the surprise. The $\beta_s$ coefficient is statistically significant when the two-day abnormal return is the dependent variable, but it loses significance when we try to explain the drift.

It could be the case that the explanatory power of the probability of information trading is driven by the medium size announcement surprises. So we re-estimate the equations only taking into account earnings announcement surprises in the top three deciles and the bottom three deciles. The results in table 10 and table 11 suggest that our results are not driven by the medium size announcements. The tables show that the PI variable has a stronger effect on the post-announcement drift when it comes to the top three surprise deciles than as a whole or when it comes to the bottom three deciles. In contrast, the earnings surprise has a stronger effect on the post-announcement drift of the bottom three deciles.

We plot in figure 7 the cumulative abnormal returns associated with a high probability of information trading (PI prior to the announcement day is above the $80^{th}$ percentile) and a low probability of information trading (PI prior to the announcement day is below the $20^{th}$ percentile). The drift is significantly different for days with high PI compared to days with low PI. On days that there is a low probability of information-based trading the drift is 71% higher than on days that there is a high probability. If we look at table 16, we can see that the post-announcement drift becomes statistically insignificant for all surprise deciles except for decile 10.

In previous studies, volume, firm size and the bid/ask spread have been able to explain the two-day abnormal return and the post-announcement drift.[16] We include these variables in equations (12), (13), (14) and (15), because they are correlated with PI, and some of PI's explanatory power could come from the correlation of PI with all of these variables. Specially the strong correlation with the firm capital size, which has been proven to be the most successful variable in explaining the post-announcement drift, could explain the statistically significance of PI. The equations are as follows,

$$|AR_{kd}| = \alpha + \beta_p PI_{kd} + \beta_s |S_{kd}| + \beta_{spread} Spread_{kd} + \beta_{vol} Vol_{kd} + \beta_c C_{kd} + \varepsilon_{kd} \qquad (16)$$

$$|CAR_{kd}| = \alpha + \beta_p PI_{kd} + \beta_s |S_{kd}| + \beta_{spread} Spread_{kd} + \beta_{vol} Vol_{kd} + \beta_c C_{kd} + \varepsilon_{kd} \qquad (17)$$

When we include the firm capital size, spread and volume the explanatory value adjusting for degrees of freedom increases, nevertheless PIBT is still statistically significant. The results are shown in table 12, table 13 and table 14. It is interesting to notice that the spread and the standardized earnings announcement surprise is not statistically significant in most instances, so only using the spread or the surprise size in order to explain the drift is not sufficient. The $\beta_p$ coefficient is always negative and statistically significant, although its magnitude is smaller. This was expected since part of the PI explanatory power comes from its correlation with firm size.

The last test we have in mind is to include in equations (16) and (17) the absolute value of past cumulative abnormal returns over forty days prior to the announcement [- 40, -1], $|CAR_{kd}|$, and the two-day abnormal return on the day of the announcement and the day after, $|AR_{kd}|$. This will allow us to see whether the explanatory power of PI is derived from either using PI as a proxy for past cumulative abnormal returns or because we are able to better measure the two-day abnormal returns. The equations are as follows,

$$|AR_{kd}| = \alpha + \beta_p PI_{kd} + \beta_s |S_{kd}| + \beta_{spread} Spread_{kd} + \beta_{vol} Vol_{kd} + \beta_c C_{kd} + \beta_{CAR} |CAR_{kd-1}| + \varepsilon_{kd} \qquad (18)$$

[16]Bernard and Thomas (1989), Maddala and Nimalendra (1995), Kandel and Pearson (1995) and Garfinkel and Sokobin (2000).

$$
\begin{aligned}
|CAR_{kd}| = \; & \alpha + \beta_p PI_{kd} + \beta_s |S_{kd}| + \beta_{spread} Spread_{kd} + \beta_{vol} Vol_{kd} \\
& + \beta_c C_{kd} + \beta_{CAR} |CAR_{kd-1}| + \beta_{AR} |AR_{kd}| + \varepsilon_{kd}
\end{aligned} \tag{19}
$$

The results are presented in table 14. These results show the importance of our measure of PI prior to the announcement. This measure is significant above and beyond what can be explained by simply including past cumulative abnormal returns and the two-day abnormal returns. What is very interesting is that we would expect a high probability of PIBT to correspond with high cumulative abnormal returns prior to the earnings announcement. Indeed PIBT is positively correlated with the absolute value of the cumulative abnormal returns prior to the earnings announcement. This correlations is 0.012. This correlation is positive but very small. This means that our PI variable prior to earnings announcements is not picking up the slow increase in abnormal returns prior to the announcement, instead it is able to detect a few "abnormal" days where there was an imbalance in the number of buys and sells. This becomes clear when we look at figure 7. This finding is interesting in itself and it deserves further attention in future research.

## 5. Conclusion

In sum, we present a new explanation for the existing gap between a theoretically efficient market and what we actually observe in the data. This explanation reconciles the empirical evidence for both small and large firms: the post-announcement drift is a function of PI in the hands of agents prior to the earnings announcements. Over time the process of trading, and learning from trading, results in prices converging to full information levels. The higher PI is the faster this convergence will be. There is no post-announcement drift when agents receive PI *before* the earnings announcement; if agents do not have PI *before* the announcement, and the public announcement is noisy, then prices need to be adjusted *after* the earnings announcement.

The most important contribution of the paper is that we are able to quantify the effect PI has on the post-announcement drift. On days that there is a low probability of information-based trading the drift is 71% higher than on days that there is a high probability. This empirical result should hold for all public announcements. We use earnings announcements because they are readily available, but in future research we should test our model with other public announcements and use different ways of measuring PIBT.

## 6. Appendix

To illustrate how this theoretical model can predict what happens after an earnings announcement we can simulate the model assuming public announcements are a noisy signal of the true value of the asset. We assume the true value of the asset follows an autoregressive process of order one with a constant term equal to c and an autoregressive coefficient equal to $\rho$,

$$V_d = c + \rho V_{d-1} + z_d \tag{A1}$$

$d$ is the day index, $d$ = 1, 2, 3 ..., $V_0$ is given, $\rho$ has an absolute value strictly less than one and $z_d$ are news innovations about the firm that affect the value of the asset; it is distributed independently and normally with mean zero and standard deviation $\sigma_z$.[17] The private information informed traders receive on day $d$ is $V_d$. This information can be summarized as a good news, a bad news or a no news signal compatible with EKOP. The good news signal means that the price set by the market maker on day $d$, time 0 labeled, $P_{d0}$ is strictly less than $V_d$. Whenever this is the case we will label the "true" value of the asset $V_d = \overline{V}_d$, as in EKOP, symbolizing that prices need to be revised up. A bad news signal means that, $P_{d0}$ is strictly greater than $V_d$. Whenever this is the case we will label $V_d = \underline{V}_d$ symbolizing that prices need to be revised down. Finally, no news, occurs with probability, $1-\alpha$, and it means that agents do not receive any private information and we will label the expected value of the asset $V_d^*$, symbolizing that prices do not need to be revised, unless there is a public announcement. On a given day we impose the following restriction $\overline{V}_d > V_d^* > \underline{V}_d$, by setting $V_d^* = (1-\delta)\overline{V}_d + \delta \underline{V}_d$.

On days $D_1, D_2, D_3, ...$ at time $T$ earnings announcements are released; these reports reveal a noisy signal of the true value of the asset,

$$\hat{V}_{D_i} = V_{D_i} + \varepsilon_{D_i},$$

$i = 1,2,3...$ where $\varepsilon_{D_i}$ is distributed independently and normally with mean zero and standard deviation $\sigma_\varepsilon$.[18] If agents receive private information on the day of the earnings announcement the

---

[17]The exogenous value of the asset is simple enough (we do not introduce a time varying volatility process) so it does not distract attention from the focus of the paper, agents price discovery, but it is rich enough to fit daily stock price behavior.

[18]In the simulation we need to assume that the public announcement is a noise signal of the true value of the asset, because otherwise the private information agents receive after the annoucement is not meaningful. In reality the private information agents receive after the announcement can either be related to the previous earnings announcement or to future announcements.

true value of the asset is attained before the announcement takes place, so $\hat{V}_{D_1}$, which is a noisy signal of the "true" value of the asset, does not provide any more information.

On the other hand, if agents do not receive private information, then the market maker uses $\hat{V}_{D_1}$ to quote a price. After the announcement is released, agents will receive private information, which will tell them whether the asset price needs further adjustment or not. The market maker's price quote will depend on the past history of earnings announcements and the private information received by informed agents up to date.[19] The market maker behaves competitively hence he earns zero profits. The zero profit condition implies that the price set by the market maker on day $D_1$ at time $T$ is equal to his expected value of the asset conditional on his information set,

$$P_{D_1T} = E(V_{D_1} | P_{1T}, P_{2T}, P_{3T}, \ldots P_{D1-1T}, Q_{D_1T}, \hat{V}_{D_1}) \tag{A2}$$

where $Q_{D_1T}$ is the trade history or the number of sells and buys the market maker has received on day $D_1$ up to time $T$. He does not condition on yesterday's order flow because he knows that informed traders receive private information once each day and the private information arrival is independent. Assuming agents have not received any private information before day $D_1$, equation (A2) has the following closed form solution,

$$P_{D_1T} = \frac{(c/(1-\rho))(1-\rho^2)\sigma_z^{-2} + \hat{V}_{D_1}\sigma_\varepsilon^{-2}}{(1-\rho^2)\sigma_z^{-2} + \sigma_\varepsilon^{-2}} \tag{A3}$$

This follows from Bayes rule Bayes rule, since $P_{D_1T} = E(V_{D_1} | \hat{V}_{D_1})$ and $z_d$ is distributed independently and normally with mean zero and standard deviation $\sigma_z$,

$$E(V_{D_1} | \hat{V}_{D_1}) = \frac{E[V_{D_1}]/Var(V_{D_1}) + \hat{V}_{D_1}/Var(\varepsilon_{D_1})}{1/Var(V_{D_1}) + 1/Var(\varepsilon_{D_1})}$$

We simulate the model forty days prior to the announcement and forty days after the announcement. We assume that the total number of days agents receive private information over this eighty day period is fixed to 40 days, fifty percent of the sample. If agents receive

[19]Since prices converge to full information levels by the end of the day, past prices set by the market maker convey the same information as the past private information agents have received.

information 36 days out of the first 40 days, then they only receive information on 4 days out of the 40 days left after the announcement. This assumption is crucial in the simulation and in the body of the paper we allow the data to speak for itself. Nevertheless, the intuition for this assumption is that if agents receive a large amount of private information before the earnings announcement, they will receive a low amount afterwards because we have attained the "true" value of the asset and there is no need to adjust prices further after the announcement.

In table A1 we show the mean absolute value of the price jump at the time of the announcement and the mean absolute value of the drift as a function of PI.

**Table A1**

Simulation: Price Jump Response

| PI prior to announcement | mean | standard deviation |
|---|---|---|
| 0.03 | 7.92 | 6.03 |
| 0.07 | 7.86 | 5.94 |
| 0.10 | 7.72 | 5.81 |
| 0.12 | 7.57 | 5.68 |
| 0.15 | 7.20 | 5.48 |
| 0.18 | 6.88 | 5.18 |
| 0.20 | 6.23 | 4.70 |
| 0.22 | 5.14 | 3.91 |
| 0.24 | 3.49 | 2.67 |

Simulation: Post-Announcement Drift Response

| PI prior to announcement | mean | standard deviation |
|---|---|---|
| 0.03 | 7.42 | 5.53 |
| 0.07 | 7.26 | 5.44 |
| 0.10 | 7.05 | 5.30 |
| 0.12 | 6.83 | 5.13 |
| 0.15 | 6.53 | 4.86 |
| 0.18 | 6.10 | 4.59 |
| 0.20 | 5.53 | 4.18 |
| 0.22 | 4.71 | 3.57 |
| 0.24 | 2.55 | 1.93 |

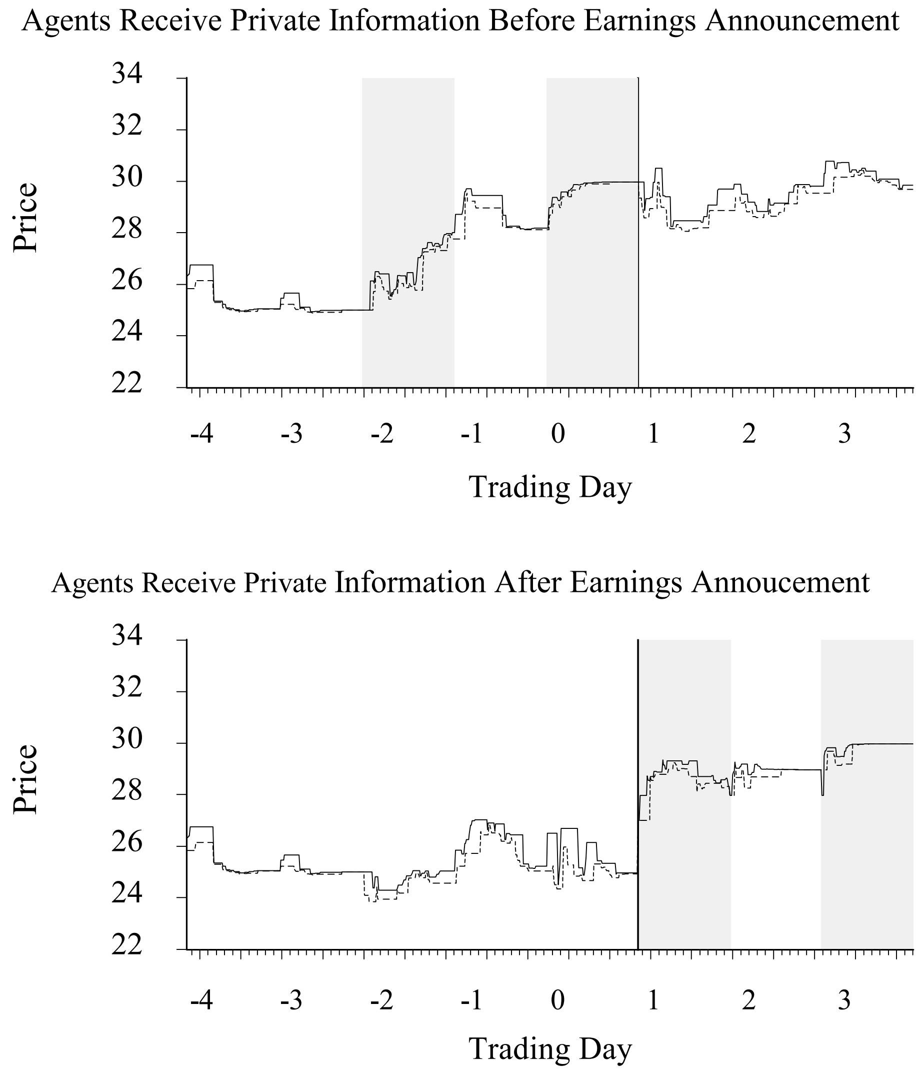
The number of simulations is 10,000. We fix the number of trades per day to be equal to 110 trades, since the NYSE is open 5 hours and 30 minutes, the trading rate is 1 transaction every three minutes. This parameter is not very important, as long as the number of trades is big enough to allow the market maker to discover the "true" value of the asset before the end of the day. The number of transactions necessary for this condition to hold varies depending on the parameters we use, but 110 transactions is enough given our parameters. We assume the true value of the asset follows an autoregressive process of order one with a constant term equal to 2.3, an autoregressive coefficient equal to 0.9, initial value $V_{0D} = \$23$, and a 20% annualized standard deviation, which are reasonable parameters for stock prices in 1993. The noise in the public announcement has a 5.20% annualized standard deviation. This parameters was calibrated to match the drift we see in the data when there is a low probability of private information trading before the earnings announcement. Like we mentioned before we need to assume that the public announcement is a noise signal of the true value of the asset, because otherwise the private information agents receive after the announcement is not meaningful. In reality the private information agents receive after the announcement can either be related to the previous earnings announcement or to future announcements. We set $\varepsilon_s = 46$, $\varepsilon_b = 50$, $\mu = 34$, $\alpha = 0.4$ and $\delta = 0.38$, corresponding to the average parameters in the first quarter of 1993. We show the results when the probability of information trading goes from 0.03 to 0.24. To allow PI, prior to an announcement, to increase, we keep all the parameters fix and increase $\alpha$ from 0.1 to 0.9 with increments of 0.1. This will allow us to obtain a range of PI from 0.03 to 0.24. The parameter $\alpha$ is equal to the fraction of days there was an information event. Given the constraint, that the total number of days agents receive private information over this eighty day period is fixed to forty, an $\alpha =$ 0.1 means that agents received private information 4 days out of the forty days prior to the earnings announcement and they will receive private information on 36 days after the earnings announcement.

Table A1 shows that the price jump at the time of the announcement and the post announcement drift are decreasing in PI. In figure A1 we show one simulation, out of the 10,000 we ran, focusing on the days surrounding the announcement. The shaded area in the figure corresponds to days when agents receive private information and the vertical solid line marks the announcement time *T*. In this example, the earnings announcement surprise is positive. The top panel shows what happens when agents receive private information prior to the earnings announcement time. There is no jump in the stock price, at the time of the announcement and there is no price adjustment afterwards, because the “true” value of the asset is already attained.

The bottom panel shows what happens if agents receive private information only after the earnings announcement and this information is positively correlated with the announcement surprise. There is a jump in the stock price equal to $P_{D_1 T-1} - P_{D_1 T}$, as defined in equation (A3), and there is a post-announcement drift given the private information agents receive after the news announcement.

**Figure A1**

Simulation: Price Dynamics and Information Trading




Notes: We plot the bid and ask quotes when informed traders receive private information before the earnings news announcement is released (top graph) and when they receive private information after the earnings news announcement (bottom graph). The solid line is the ask quote and the dashed line is the bid quote. The vertical line shows the time when the earnings announcement was released. The shaded area highlights the days informed traders received private information. In the top graph there is no price jump at the time of the news announcement and there is no post-announcement drift. In the bottom graph there is a price jump at the time of the earnings announcement (7.45 % price increase) and there is a post-announcement drift of (5.37 %).

## References


Admati, Anat R. and Paul Pfleiderer (1988), "A Theory of Intraday Patterns: Volume and Price Variability," *Review of Financial Studies*, 1, 3-40.

Ball, R. and P. Brown (1968), "An Empirical Evaluation of Accounting Income Numbers," *Journal of Accounting Research* 159-78.

Barberis, Nicholas, Andrei Shleifer and Robert Vishny (1998), "A Model of Investor Sentiment," *Journal of Financial Economics,* 49, 307-343.

Bernard, Victor L. and Jacob K. Thomas (1989), "Post-Earnings-Announcement Drift: Delayed Price Response or Risk Premium?," *Journal of Accounting Research*, 27, 1-36.

Copeland, T. and D. Galai (1983), "Information Effects on the Bid-Ask Spread," *Journal of Finance,* 38, 1457-1469.

Cutler, David M., James M. Poterba and Lawrence H. Summers (1990), "Speculative Dynamics," *Review of Economic Studies,* 58, 529-546.

De Long, J. B., Shleifer, A., Summers, L. and Waldmann, R. (1990), "Noise Trader Risk in Financial Markets," *Journal of Political Economy* 98, 703-738.

Diamond, W. Douglas and Robert E. Verrecchia (1991), "Liquidity and the Cost of Capital," *Journal of Finance,* 46, 1325-1359.

Easley, David, Nicholas M. Kiefer, Maureen O'Hara and Joseph B. Paperman (1996), "Liquidity, Information, and Infrequently Traded Stocks," *Journal of Finance,* 51, 1405-1437.

Easley, David, Soeren Hvidkjaer and Maureen O'Hara (2001), "Is Information Risk a Determinant of Asset Returns?" *working paper,* Cornell University.

Fama, Eugene F. (1970), "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance*, 25, 383-417.

Fama, Eugene F. and Kenneth R. French (1996), "Multifactor Explanations of Asset Pricing Anomalies," *Journal of Finance,* 51, 55-84.

Fama, Eugene F. (1998), "Market Efficiency, Long-Term Returns, and Behavioral Finance," *Journal of Financial Economics,* 49, 283-306.

Hasbrouck, Joel (1991), "Measuring the Information Content of Stock Trades," *Journal of Finance*, 46, 179-207.

Garfinkel, Jon and Jonathan Sokobin (2000), "Rational Markets, Earnings Transparency and Post-Earnings Announcement Drift," *working paper,* University of Iowa.

Goldberger, A.S. (1972), "Maximum-likelihood Estimation of Regression Models Containing Unobservable Variables," *International Economic Review*, 13, 1-15.

Glosten, Lawrence R., and Paul R. Milgrom (1985), "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders," *Journal of Financial Economics* 14, 71-100.

Grilliches, Z. (1974), "Errors in Variables and Other Unobservables," *Econometrica,* 42, 971-998.

Hong, Harrison and Jeremy C. Stein (1999), "A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets," *Journal of Finance,* 54, 2143-2184.

Kandel, Eugene and Neil D. Pearson (1995), "Differential Interpretation of Public Signals and Trade in Speculative Markets," *Journal of Political Economy,* 103, 831-872.

Kyle, Albert S. (1985), "Continuous Auctions and Insider Trading," *Econometrica*, 53, 1315-1335.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny (1994), "Contrarian Investment, Extrapolation, and Risk," *Journal of Finance,* 49, 1541-1578.

Lee, Charles M. and Mark Ready (1991), "Inferring Trade Direction from Intraday Data," *Journal of Finance,* 46, 733-746.

Maddala, G.S. and M. Nimalendran (1995), "An Unobserved Component Panel Data Model to Study the Effect of Earnings Surprise on Stock Prices, Trading Volumes and Spreads," *Journal of Econometrics,* 68, 229-242.

Murphy, Kevin and Robert Topel (1985), "Estimation and Inference in Two-Step Econometric Models," *Journal of Business and Economic Statistics,* 3, 370-379.

Toe, S., I. Welch and T. Wong (1998), "Earnings Management and the Long-run Market Performance of Initial Public Offerings," *Journal of Finance,* 53, 1935-1974.

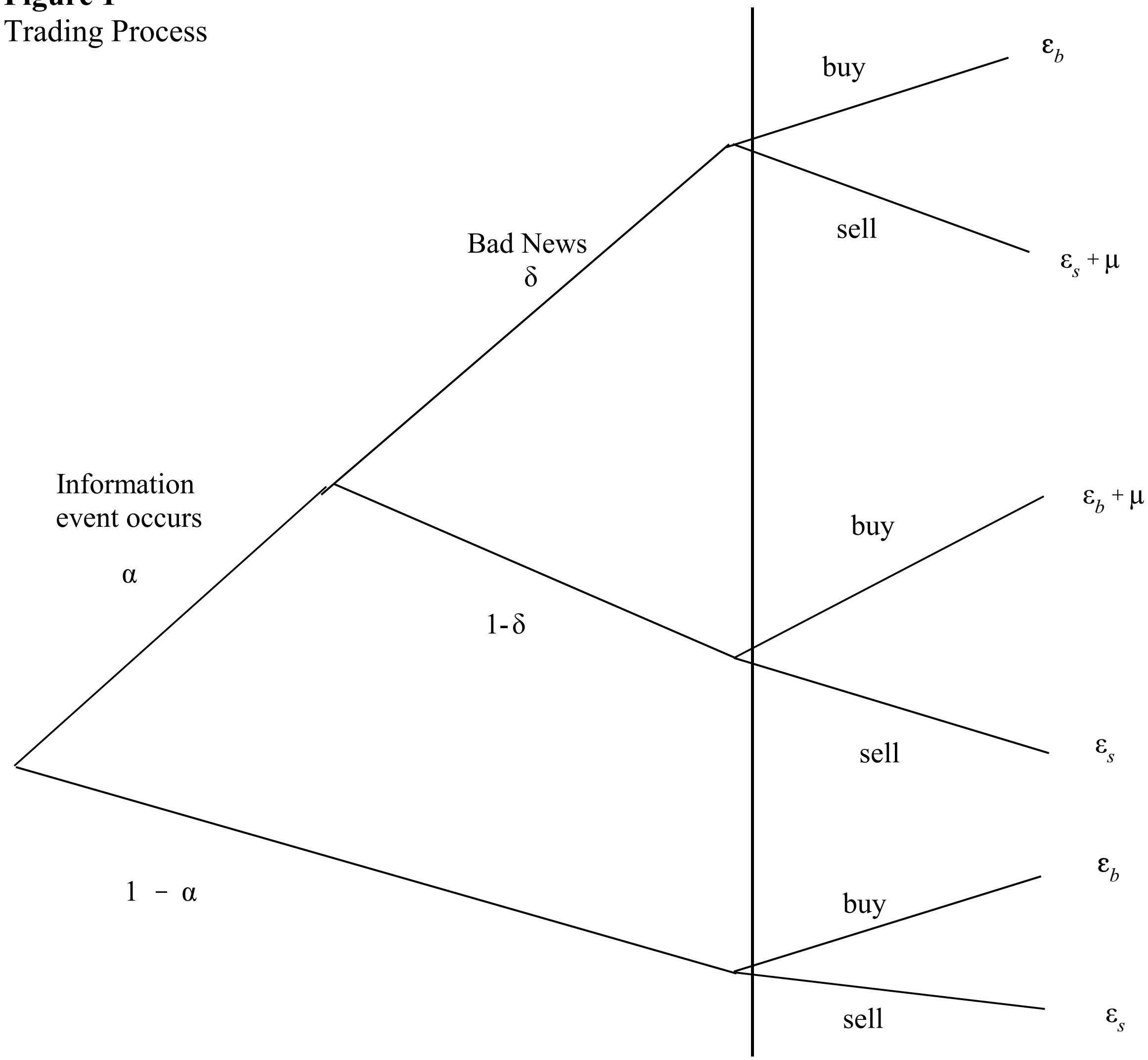
Sheski, David (1997), "Handbook of Parametric and Nonparametric Statistical Procedures," CRC Press.

Shleifer, Andrei and Vishny, Robert (1997), "The Limits of Arbitrage," *Journal of Finance,* 52, 35-55.

Zarowin, Paul (1989), "Does the Stock Market Overreact to Corporate Earnings Information?," *Journal of Finance,* 44, 1385-1400.

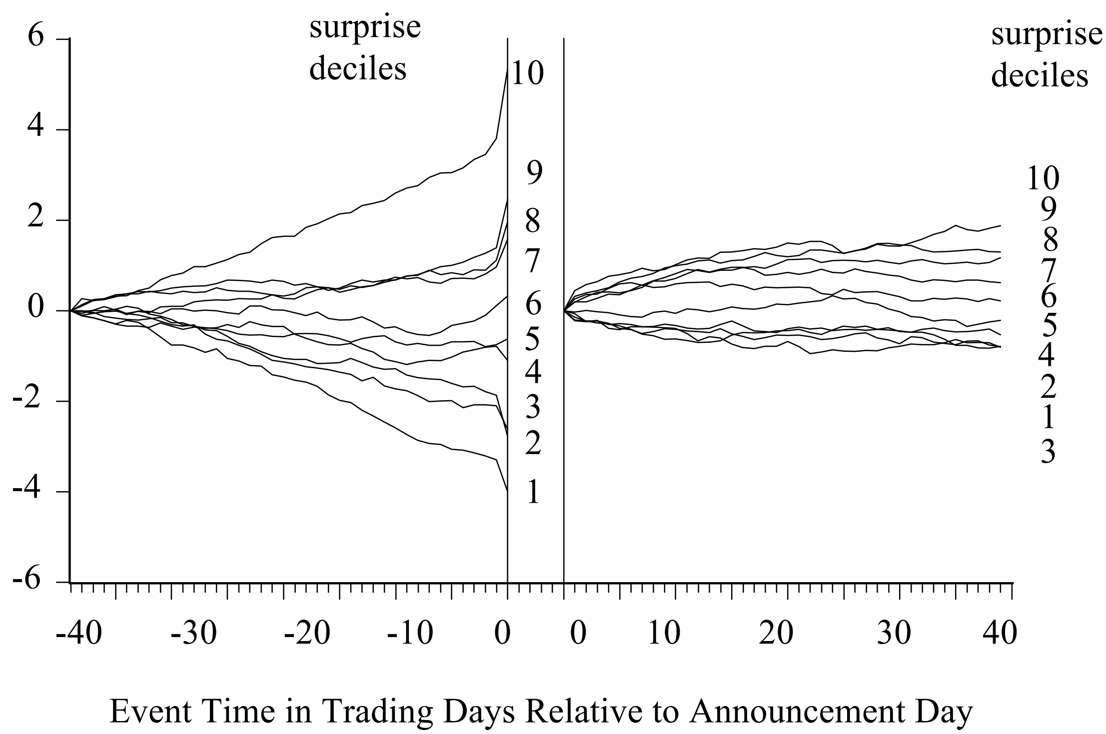
Zellner, A. (1970), "Estimation of Regression Relationships Containing Unobservable Independent Variables," *International Economic Review,* 11, 441-454.

**Figure 1**
Trading Process



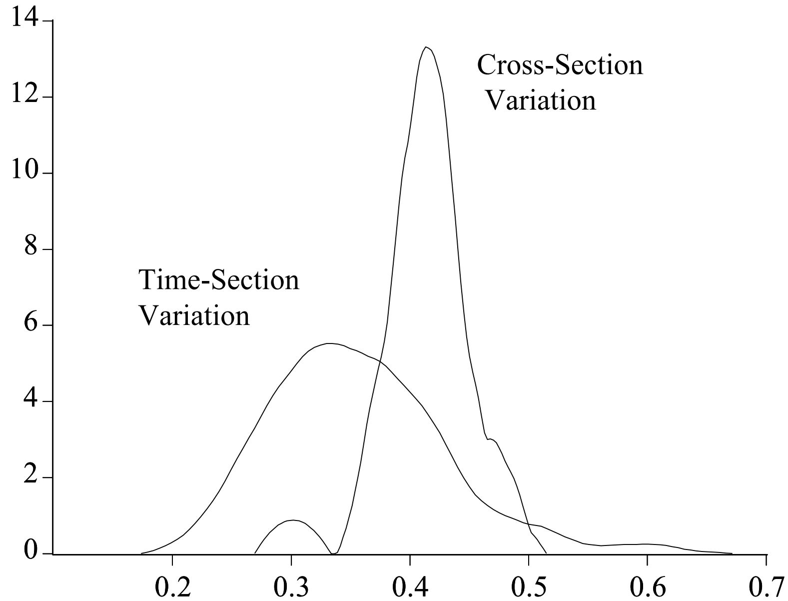

Notes: This is a tree diagram of the trading process taken from EKOP (1996). $\alpha$ is the probability of an information event, $\delta$ is the probability of a low signal, $\mu$ is the rate of informed trade arrival, $\varepsilon_b$ is the uninformed buy trade arrival and $\varepsilon_s$ is the uninformed sell trade arrival. Nodes to the left of the solid vertical line occur once a day.

**Figure 2**
Cumulative Abnormal Returns 1993-1999




Notes: We plot the means of cumulative size adjusted returns over trading days [- 40, 0] before the earnings announcement date, day 0, and over trading days [+1, +40]. The returns are categorized by surprise deciles. The surprise decile is a ranking from 1 to 10 of this quarter's surprise $S_{kt}$ (firm k, quarter t), based on the previous quarter's surprise decile cut offs. See text for details.

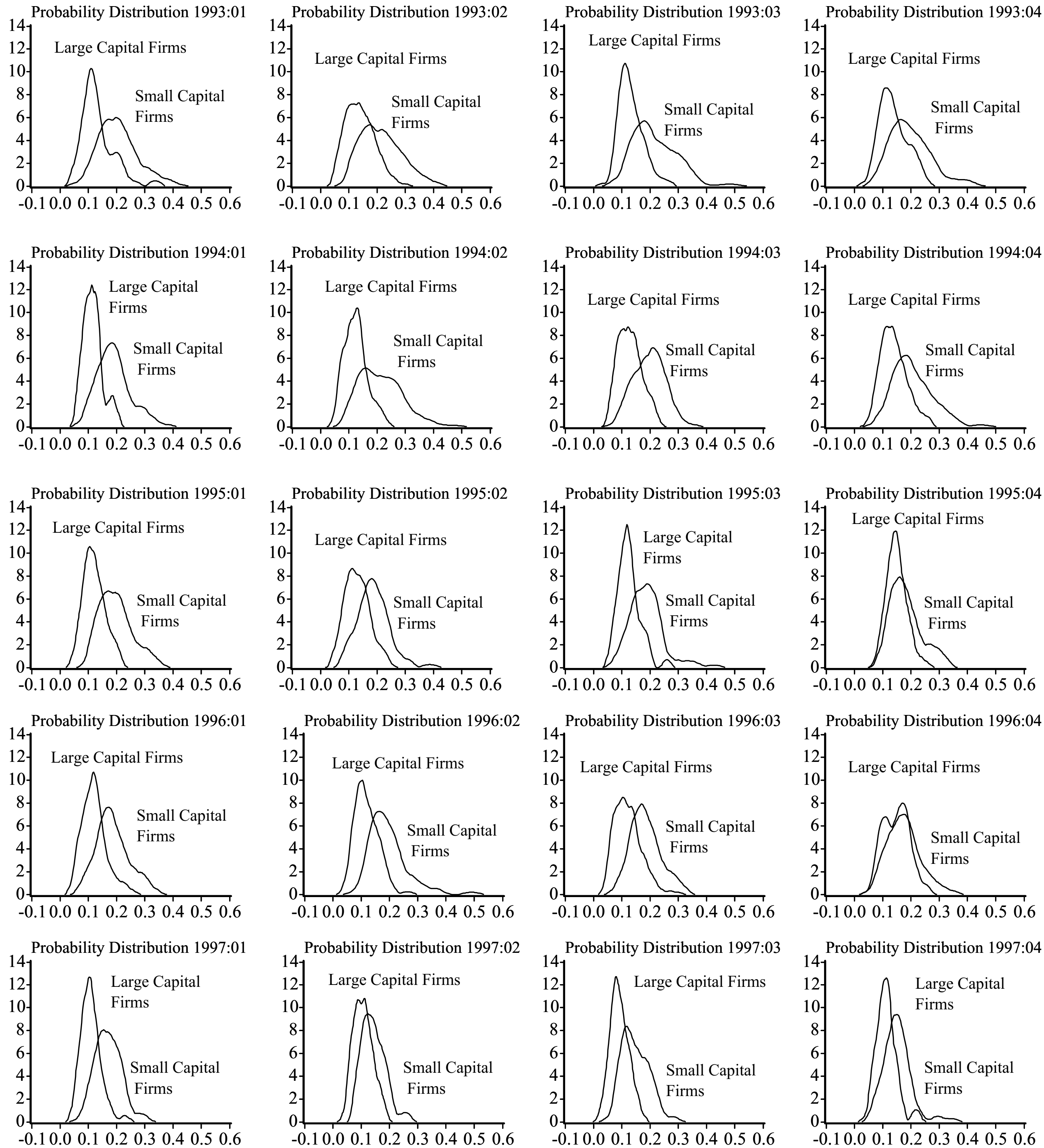
**Figure 3**
Private Information Coefficient of Variation




| | Time-Series | Cross-Section |
|---|---|---|
| Mean | 0.35 | 0.41 |
| Median | 0.34 | 0.41 |
| Max | 0.62 | 0.48 |
| Min | 0.22 | 0.30 |
| Std. Deviation | 0.07 | 0.03 |
| Skewness | 0.78 | -0.75 |
| Kurtosis | 3.98 | 5.1 |

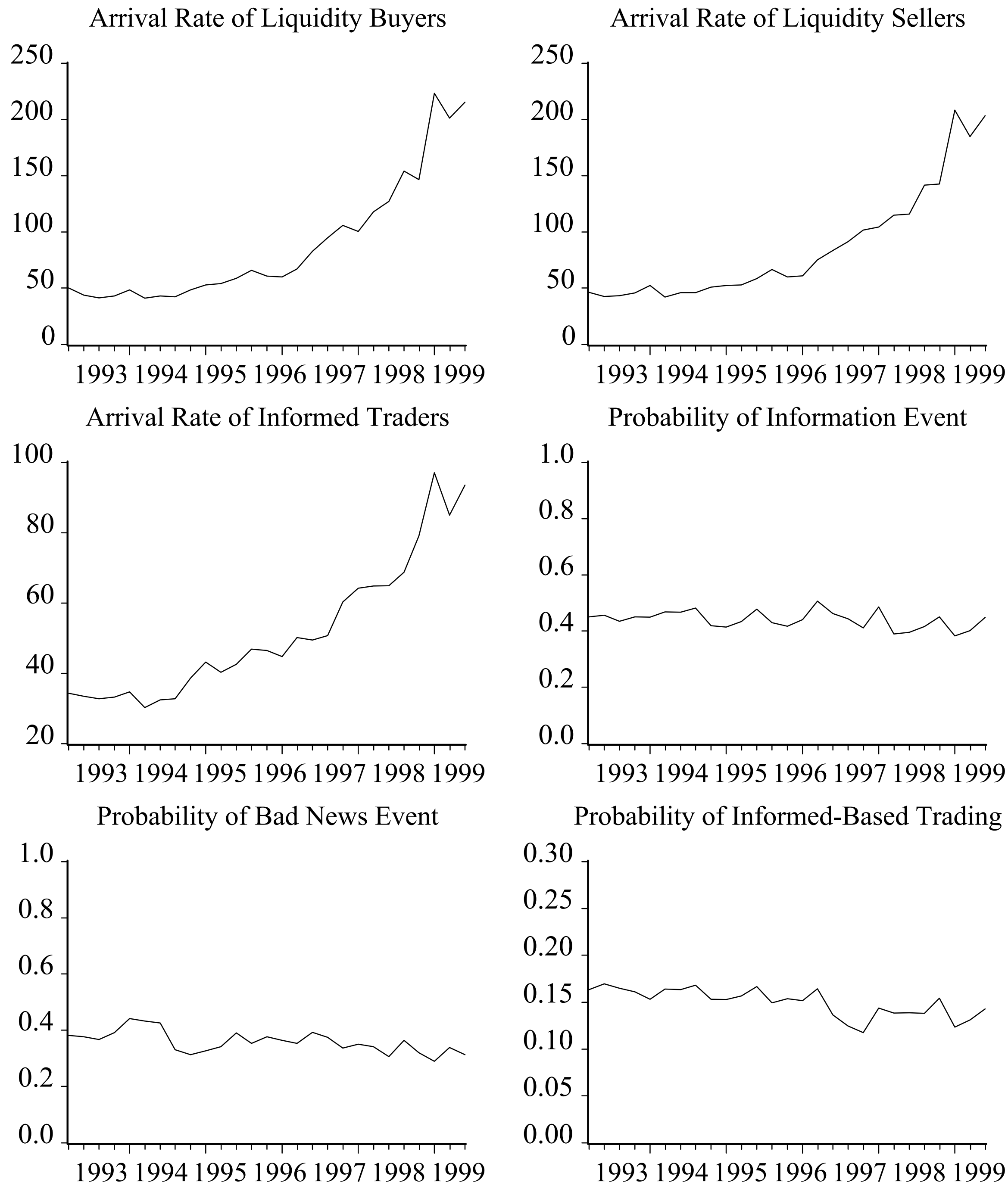
Notes: We show the kernel estimates of the density of the coefficient of variation of PI across time and across stocks. We use the Epanechnikov kernel as a weighting function and we use Silverman's bandwidth selection criteria. The coefficient of variation across time is calculated by fixing a stock and dividing the standard deviation of PI across time by the mean of PI across time. To calculate the coefficient of variation across stocks we fix a time period and divide the standard deviation of PI across stocks by the mean of PI across stocks. We test the hypothesis that the two samples have the same median against the alternative hypothesis that they have different medians. We obtain a Mann-Whitney statistic of 4.8; since the critical value at the five percent confidence level is 1.96, we reject the null hypothesis in favor of the alternative.

**Figure 4**
Private Information Cross-Sectional Probability Distribution




Notes: We plot the kernel estimates of the density of PI for "large" firms and "small" firms from 1993:01 to 1997:04. See text for details.

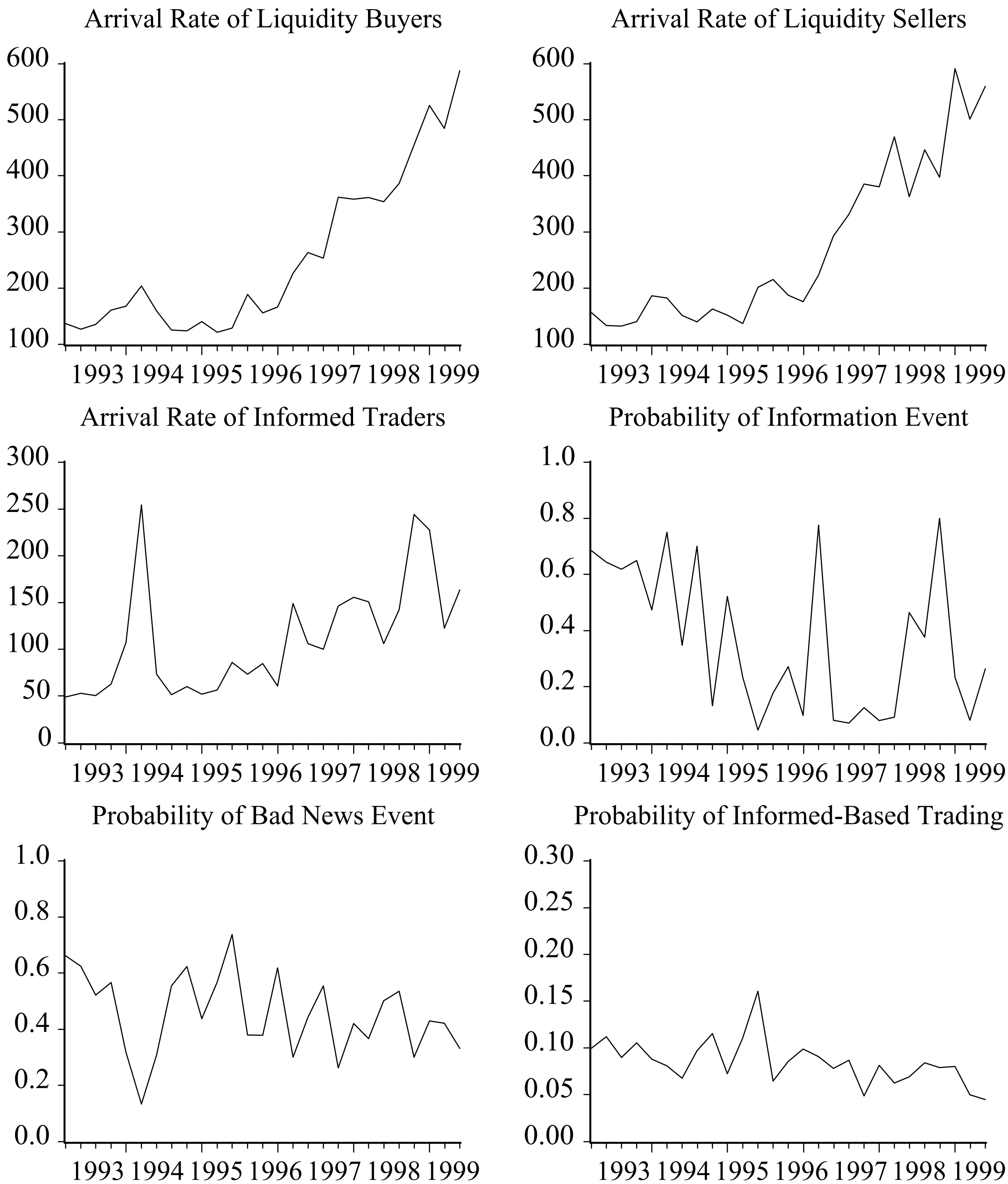
**Figure 5**
Information Trading Parameter Estimates Before an Earnings Announcement




Notes: We plot the microstructure theory parameter's average across 746 stocks. The arrival rate of liquidity sellers, $\varepsilon_s$, the arrival rate of liquidity buyers, $\varepsilon_b$, the arrival rate of informed traders, $\mu$, the probability of an information event, $\alpha$, the probability of bad news, $\delta$, and the probability of information trading, PI, are estimated using the likelihood function given in equation (1). The time series plot is obtained by calculating each parameter using data covering the 40 days prior to an earnings announcement day and averaging across stocks. On average each stock has 23 estimates.

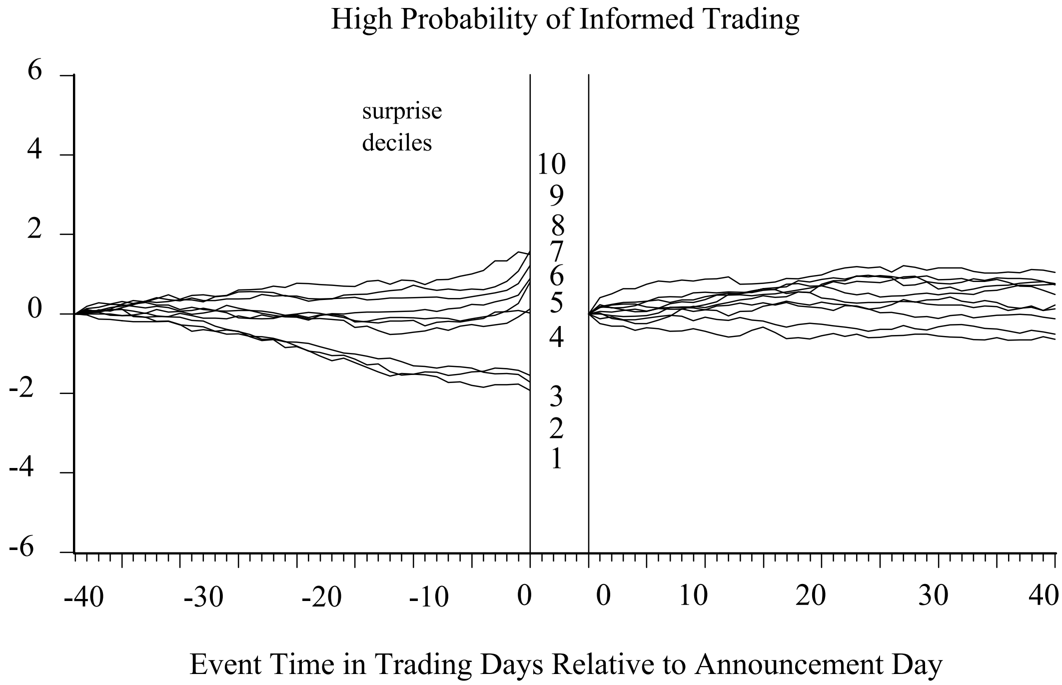
**Figure 6**

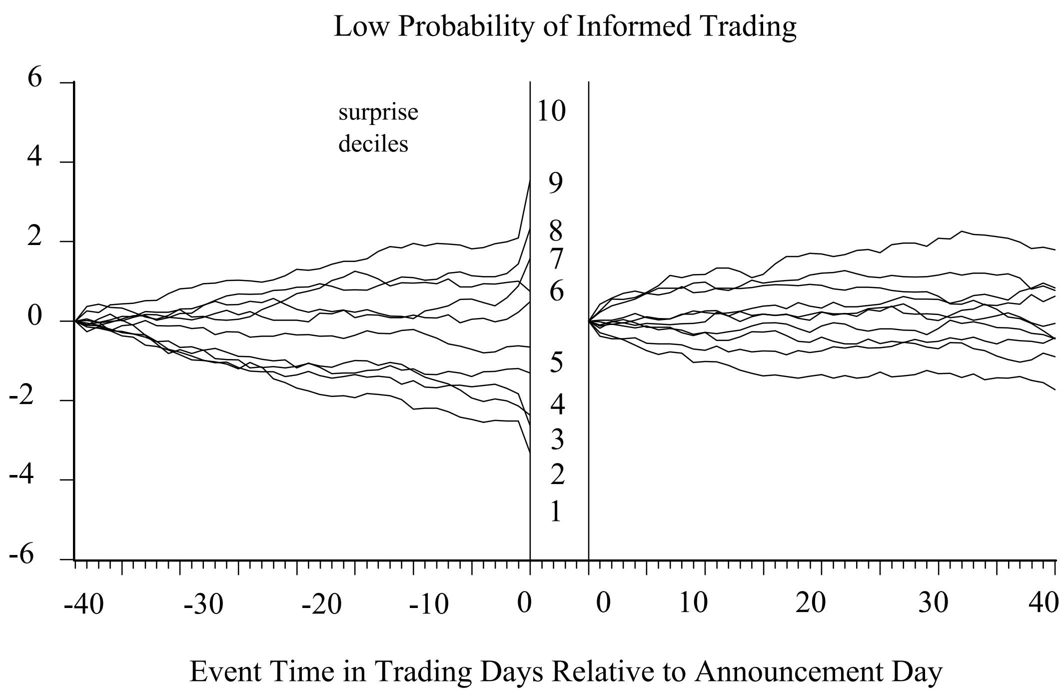
Exxon's Information Trading Parameter Estimates Before an Earnings Announcement




Notes: We plot the microstructure theory parameters for Exxon Corporation. The arrival rate of liquidity sellers, $\varepsilon_s$, the arrival rate of liquidity buyers, $\varepsilon_b$, the arrival rate of informed traders, $\mu$, the probability of an information event, $\alpha$, the probability of bad news, $\delta$, and the probability of information trading, PI, are estimated using the likelihood function given in equation (1). The time series plot is obtained by calculating each parameter using data covering the 40 days prior to an earnings announcement date. In total we have 27 estimates.

**Figure 7**




Notes: We plot the cumulative size adjusted returns over trading days [- 40, 0] before the earnings announcement date, day 0, and over trading days [+1, +40]. The high probability (low probability) of informed trading is defined as days when the probability of information trading prior to the earnings announcement was above the 80th percentile (below the 20th percentile). On days that there is a low probability of information-based trading the drift is 71% higher than on days that there is a high probability.

**Table 1**
Parameter Simulation

| | $\alpha$ | $\delta$ | $\mu$ | $\varepsilon_b$ | $\varepsilon_s$ | PI |
|---|---|---|---|---|---|---|
| Both News | 0.40 | 0.506 | 19.333 | 39.630 | 39.908 | 0.102 |
| Bad News | 0.40 | 1.000 | 18.106 | 40.098 | 40.653 | 0.094 |
| Good News | 0.40 | 0.000 | 18.985 | 40.301 | 40.000 | 0.101 |

Notes: We generate three samples of forty days, where buys and sells are generated as normally, independently and identically distributed variables with mean equal to forty and standard deviation equal to three. In the first sample labeled "good news" we choose at random sixteen days out of the forty days and we add twenty buys more to the "usual" level of buys. The "bad news" sample consists of sixteen days where there are twenty sells more than "usual." Finally, in the sample "both news" we choose at random sixteen days out of the forty days and we add twenty buys more to the "usual" level of buys, we choose again at random sixteen days and we add twenty sells more to the "usual" level of buys. We calculate the likelihood function using these three different samples.

**Table 2**
Post-Announcement Drift Stylized Facts

| Decile | Standardized Surprise | 2-Day Ab. Returns % | Drift % |
|---|---|---|---|
| 1 | -1.9056 | -0.99* | -0.92* |
| 2 | -0.7619 | -0.78* | -0.89* |
| 3 | -0.3867 | -0.54 | -0.98* |
| 4 | -0.1314 | -0.29 | -0.48 |
| 5 | 0.0288 | 0.17 | -0.51 |
| 6 | 0.1900 | 0.43 | 0.15 |
| 7 | 0.4099 | 0.77 | 0.60 |
| 8 | 0.6593 | 1.07* | 1.20* |
| 9 | 1.0571 | 1.19* | 1.24* |
| 10 | 2.0051 | 1.89* | 2.02* |

Notes: We present in this table the means of variables, categorized by surprise deciles. The surprise decile is a ranking from 1 to 10 of this quarter's surprise $S_{kt}$ (firm k, quarter t), based on the previous quarter's surprise decile cutoffs. The two-day abnormal return is the cumulative size adjusted return on the day of the announcement and the day after. We look at the day after, because some earnings announcements take place after the stock market is closed and the price reaction is observed the next day. Drift is calculated as the cumulated size adjusted return over trading days [+1,+40], where day 0 is the day after the earnings announcement date. The asterisk denotes that the abnormal returns are statistically significant at the five percent level.

**Table 3**

Post-Announcement Drift Stylized Facts: Small and Large Firms

| | Small Capital Firms | | Large Capital Firms | |
|---|---|---|---|---|
| Decile | 2-Day Ab. Returns % | Drift % | 2-Day Ab. Returns % | Drift % |
| 1 | -0.89* | -1.24* | -0.96* | -1.05* |
| 2 | -0.76* | -1.02* | -0.65 | -1.30* |
| 3 | -0.85* | -0.62 | -0.28 | -0.53 |
| 4 | -0.19 | -0.24 | -0.19 | -0.45 |
| 5 | 0.04 | 0.22 | -0.17 | -0.26 |
| 6 | 0.49 | 0.82* | 0.05 | 0.24 |
| 7 | 0.78 | 1.10* | 0.30 | 0.46 |
| 8 | 0.92* | 1.97* | 0.72 | 0.93* |
| 9 | 1.39* | 2.85* | 0.95* | 0.96* |
| 10 | 1.77* | 3.20* | 1.55* | 1.05* |

Notes: We present in this table the means of variables, categorized by surprise deciles. The surprise decile is a ranking from 1 to 10 of this quarter's surprise $S_{kt}$ (firm k, quarter t), based on the previous quarter's surprise decile cutoffs. The two-day abnormal return is the cumulative size adjusted return on the day of the announcement and the day after. We look at the day after, because some earnings announcements take place after the stock market is closed and the price reaction is observed the next day. Drift is calculated as the cumulated size adjusted return over trading days [+1,+40], where day 0 is the day after the earnings announcement date. The asterisk denotes that the abnormal returns are statistically significant at the five percent level.

**Table 4**
Simple Correlations

| | PI | Standardized surprise | 2-Day Ab. Returns | Drift | First half of the | Second half of the | Market Cap. Size | Volume |
|---|---|---|---|---|---|---|---|---|
| Standardized surprise | -0.0639 | | | | | | | |
| 2-Day Ab. Returns | -0.026 | 0.1012 | | | | | | |
| Drift | -0.0351 | 0.0121 | 0.1398 | | | | | |
| First half of the drift | -0.0049 | 0.0019 | 0.1542 | 0.5665 | | | | |
| Second half of the | -0.0384 | 0.0147 | 0.1131 | 0.5866 | 0.1573 | | | |
| Market Capital Size | -0.5092 | 0.051 | -0.0906 | -0.1402 | -0.1701 | -0.1266 | | |
| Volume | -0.5875 | 0.0885 | 0.1518 | 0.0778 | 0.0538 | 0.0879 | 0.6292 | |
| Spread | 0.3462 | -0.0559 | -0.0684 | -0.0018 | 0.0012 | -0.0067 | -0.3375 | -0.2841 |

Notes: We show a simple correlation matrix of $PI_{kd}$ against the absolute value of the standardized earnings announcement surprise, $|S_{kd}|$, the absolute value of the two-day abnormal returns, $|AR_{kd}|$, the absolute value of the post-announcement cumulative abnormal returns over forty days, $|CAR_{kd}|$, the first 20 days [+1, +20], $|CAR_{1kd}|$, and the last 20 days [+21, +40], $|CAR_{2kd}|$, the market value capital decile, $C_{kd}$ , a cardinal number from 1 to 10 corresponding to small, medium and large firms, spread, which is calculated as the average daily spread in the forty days prior to the earnings announcement, volume, which is the logarithm of the average daily dollar trading volume in the forty days prior to the announcement. See text for more details.

**Table 5**
Mann-Whitney Tests for Equality of Medians Between the Large Firm's Probability of Information Trading and Small Firm's Probability of Information Trading

| Year | Quarter | Mann-Whitney Test |
|---|---|---|
| 1993 | 1 | 7.55 |
| | 2 | 7.07 |
| | 3 | 7.94 |
| | 4 | 6.19 |
| 1994 | 1 | 8.71 |
| | 2 | 8.66 |
| | 3 | 7.90 |
| | 4 | 6.95 |
| 1995 | 1 | 8.57 |
| | 2 | 6.59 |
| | 3 | 7.64 |
| | 4 | 4.03 |
| 1996 | 1 | 7.76 |
| | 2 | 8.49 |
| | 3 | 7.57 |
| | 4 | 3.07 |
| 1997 | 1 | 8.43 |
| | 2 | 5.81 |
| | 3 | 7.80 |
| | 4 | 5.64 |
| 1998 | 1 | 6.84 |
| | 2 | 7.01 |
| | 3 | 7.70 |
| | 4 | 4.41 |
| 1999 | 1 | 4.39 |
| | 2 | 5.20 |
| | 3 | 4.43 |

Notes: The Mann-Whitney statistic is approximately normally distributed. We test the hypothesis that the two samples have the same median, against the alternative hypothesis that they have different medians. The critical value at the five percent confidence level is 1.96, so we reject the null hypothesis in favor of the alternative for all the periods.

**Table 6**

| Year | Quarter | Likelihood Ratio Test |
|---|---|---|
| 1993 | 1 | 30.57% |
| | 2 | 10.48% |
| | 3 | 6.55% |
| | 4 | 1.75% |
| 1994 | 1 | 1.31% |
| | 2 | 0.87% |

Notes: We calculate the constrained maximum likelihood estimator, $\hat{\theta}_c = \{\hat{\varepsilon}_b, \hat{\varepsilon}_s, \hat{\mu}, \hat{\delta}, \hat{\alpha}\}$, for each stock by maximizing equation (1) constraining the parameters to be the same across time. We also calculate the unconstrained maximum likelihood estimator, $\hat{\theta}_u$, for each stock by maximizing equation (1) allowing the parameters to vary over time. To test the hypothesis we further restrict ourselves to the sample period from 1993 to 1994 where we reject the null hypothesis of parameters $\varepsilon_b, \varepsilon_s,$ and $\mu$ trending for 70% of the stocks. We test the null hypothesis $\theta = \hat{\theta}_c$ using a likelihood ratio test. The likelihood ratio test statistic is equal to $-2(\ln\hat{L}_c - \ln\hat{L}_u)$, where $\hat{L}_c$ is the likelihood function described in equation 1 evaluated at $\hat{\theta}_c$ and $\hat{L}_u$ is the same likelihood function evaluated at $\hat{\theta}_u$. The large sample distribution of this test statistic is a chi-square distribution with the number of degrees of freedom equal to the number of restrictions. We present the fraction of stocks for which we accept the null hypothesis when we constrain our parameters over two quarters, three quarters, four quarters, five quarters and six quarters.

**Table 7**
Time-Series Analysis

| | $\beta_p$ | $\beta_s$ |
|---|---|---|
| $\lvert AR_d \rvert = \alpha + \beta_p PI_d + \varepsilon_d$ | | |
| Statistically Significant | 13.10% | |
| Negative Coefficient | 60.70% | |
| Statistically Significant Negative Coefficient | 12.48% | |
| $\lvert AR_d \rvert = \alpha + \beta_s \lvert S_d \rvert + \varepsilon_d$ | | |
| Statistically Significant | | 14.85% |
| Positive Coefficient | | 64.63% |
| Statistically Significant Positive Coefficient | | 13.97% |
| $\lvert AR_d \rvert = \alpha + \beta_p PI_d + \beta_s \lvert S_d \rvert + \varepsilon_d$ | | |
| Statistically Significant | 12.23% | 16.16% |
| Negative/Positive Coefficient | 57.64% | 63.32% |
| Stat. Significant Negative/Positive Coefficient | 11.86% | 13.10% |
| $\lvert CAR_d \rvert = \alpha + \beta_p PI_d + \varepsilon_d$ | | |
| Statistically Significant | 13.10% | |
| Negative Coefficient | 72.93% | |
| Statistically Significant Negative Coefficient | 12.23% | |
| $\lvert CAR_d \rvert = \alpha + \beta_s \lvert S_d \rvert + \varepsilon_d$ | | |
| Statistically Significant | 11.35% | |
| Positive Coefficient | 56.29% | |
| Statistically Significant Positive Coefficient | 8.24% | |
| $\lvert CAR_d \rvert = \alpha + \beta_p PI_d + \beta_s \lvert S_d \rvert + \varepsilon_d$ | | |
| Statistically Significant | 12.66% | 11.35% |
| Negative/Positive Coefficient | 73.36% | 56.29% |
| Stat. Significant Negative/Positive Coefficient | 11.79% | 8.24% |

Notes: We show the percentage number of stocks whose $\beta_p$ and $\beta_s$ coefficients in the different equations is statistically significant. Each time-series regressions have from 27 to 20 observations, and we run the regressions for 656 stocks. We use a bootstrap method to derive the variance-covariance estimator.

**Table 8**
Cross-Section Analysis

| | $\beta_p$ | $\beta_s$ |
|---|---|---|
| $\|AR_d\| = \alpha + \beta_p PI_d + \varepsilon_d$ | | |
| Statistically Significant | 33.33% | |
| Negative Coefficient | 81.48% | |
| Statistically Significant Negative Coefficient | 33.33% | |
| $\|AR_d\| = \alpha + \beta_s \|S_d\| + \varepsilon_d$ | | |
| Statistically Significant | | 33.33% |
| Positive Coefficient | | 96.30% |
| Statistically Significant Positive Coefficient | | 33.33% |
| $\|AR_d\| = \alpha + \beta_p PI_d + \beta_s \|S_d\| + \varepsilon_d$ | | |
| Statistically Significant | 29.63% | 37.04% |
| Negative/Positive Coefficient | 81.48% | 96.30% |
| Stat. Significant Negative/Positive Coefficient | 29.63% | 37.04% |
| $\|CAR_d\| = \alpha + \beta_p PI_d + \varepsilon_d$ | | |
| Statistically Significant | 24.81% | |
| Negative Coefficient | 62.96% | |
| Statistically Significant Negative Coefficient | 24.81% | |
| $\|CAR_d\| = \alpha + \beta_s \|S_d\| + \varepsilon_d$ | | |
| Statistically Significant | | 9.70% |
| Positive Coefficient | | 54.44% |
| Statistically Significant Positive Coefficient | | 9.70% |
| $\|CAR_d\| = \alpha + \beta_p PI_d + \beta_s \|S_d\| + \varepsilon_d$ | | |
| Statistically Significant | 24.81% | 7.11% |
| Negative/Positive Coefficient | 62.96% | 57.04% |
| Stat. Significant Negative/Positive Coefficient | 24.81% | 7.11% |

Notes: We show the percentage number of stocks whose $\beta_p$ and $\beta_s$ coefficients in the different equations is statistically significant. The cross-section regressions have from 746 to 540 observations. We run each regression for all 27 periods from the first quarter of 1993 to the third quarter of 1999. We use a bootstrap method to derive the variance-covariance estimator.

**Table 9**

Pooled Data Analysis: PIBT and the Standardized Surprise Response Coefficients

| | $\beta_p$ | $\beta_s$ |
|---|---|---|
| Two-Day Abnormal Return | | |
| | -0.0401* | |
| | | 0.0033* |
| | -0.0367* | 0.0031* |
| Drift | | |
| | -0.1370* | |
| | | 0.0010* |
| | -0.1368* | 0.0002 |
| First half of the drift | | |
| | -0.0800* | |
| | | 0.0013* |
| | -0.0790* | 0.0009 |
| Second half of the drift | | |
| | -0.1030* | |
| | | 0.0005 |
| | -0.1030* | 0.0001 |

Notes: We estimate the two-day abnormal returns model, $|AR_{kd}| = \alpha + \beta_p PI_{kd} + \beta_s |S_{kd}| + \varepsilon_{kd}$, and the cumulative abnormal return model, $CAR_{kd}| = \alpha + \beta_p PI_{kd} + \beta_s |S_{kd}| + \varepsilon_{kd}$ over forty days after the announcement, $|CAR_{kd}|$, over the first twenty days after the earnings announcement. We report the $\hat{\beta}_p$ and $\hat{\beta}_s$ values, and we mark with an asterisk those coefficients that are statistically significant at the five percent level, using bootstrap standard errors. We ran a weighted least squares regression to control for the heteroskedasticity in the data since $\mathrm{Var}(\varepsilon_{kd}) \propto \frac{1}{C_{kd}^2}$, where $C_{kd}$ = market capital

**Table 10**
Pooled Data Analysis: Top Three Deciles

| | $\beta_p$ | $\beta_s$ |
|---|---|---|
| Two-Day Abnormal Return | | |
| | -0.0432* | |
| | | 0.0046* |
| | -0.0365* | 0.0043* |
| Drift | | |
| | -0.1704* | |
| | | 0.0010* |
| | -0.1724* | 0.0001 |
| First half of the drift | | |
| | -0.0659* | |
| | | 0.0007* |
| | -0.0656* | 0.0001 |
| Second half of the drift | | |
| | -0.1408* | |
| | | 0.0015* |
| | -0.1403* | 0.0003 |

Notes: We repeat the same exercise as before only using observations in the top three surprise deciles. We estimate the two-day abnormal returns model, $|AR_{kd}| = \alpha + \beta_p PI_{kd} + \beta_s |S_{kd}| + \varepsilon_{kd}$, and the cumulative abnormal return model, $CAR_{kd}| = \alpha + \beta_p PI_{kd} + \beta_s |S_{kd}| + \varepsilon_{kd}$ over forty days after the announcement, $|CAR_{kd}|$, over the first twenty days after the earnings announcement, $|CAR_{1kd}|$, and over the last twenty days after the earnings announcement, $|CAR_{2kd}|$. We report the $\hat{\beta}_p$ and $\hat{\beta}_s$ values, and we mark with an asterisk those coefficients that are statistically significant at the five percent level, using bootstrap standard errors. We ran a weighted least squares regression to control for the heteroskedasticity in the data since $\mathrm{Var}(\varepsilon_{kd}) \propto \frac{1}{C_{kd}^2}$, where $C_{kd}$ = market capital value.

**Table 11**

Pooled Data Analysis: Bottom Three Deciles

| | $\beta_p$ | $\beta_s$ |
|---|---|---|
| Two-Day Abnormal Return | | |
| | -0.0308* | |
| | | 0.0020* |
| | -0.0286* | 0.0019* |
| | | |
| Drift | -0.1346* | |
| | | 0.0040* |
| | -0.1307* | 0.0035* |
| First half of the drift | | |
| | -0.0846* | |
| | | 0.0028* |
| | -0.0819* | 0.0025* |
| Second half of the drift | | |
| | -0.1079* | |
| | | 0.0026* |
| | -0.1055* | 0.0021 |

Notes: We repeat the same exercise as before only using observations in the top three surprise deciles. We estimate the two-day abnormal returns model, $|AR_{kd}| = \alpha + \beta_p PI_{kd} + \beta_s |S_{kd}| + \varepsilon_{kd}$, and the cumulative abnormal return model, $|CAR_{kd}| = \alpha + \beta_p PI_{kd} + \beta_s |S_{kd}| + \varepsilon_{kd}$, over forty days after the announcement, $|CAR_{kd}|$, over the first twenty days after the earnings announcement, $|CAR_{1kd}|$, and over the last twenty days after the earnings announcement, $|CAR_{2kd}|$. We report the $\hat{\beta}_p$ and $\hat{\beta}_s$ values, and we mark with an asterisk those coefficients that are statistically significant at the five percent level, using bootstrap standard errors. We ran a weighted least squares regression to control for the heteroskedasticity in the data since $\mathrm{Var}(\varepsilon_{kd}) \propto \frac{1}{C_{kd}^2}$, where $C_{kd}$ = market capital

**Table 12**

Pooled Data Analysis: Firm Size, Volume and Spread Effects

| | $\beta_p$ | $\beta_s$ | $\beta_{spread}$ | $\beta_{vol}$ | $\beta_c$ |
|---|---|---|---|---|---|
| Two-Day Abnormal Return | | | | | |
| | -0.0258* | 0.0030* | -0.0085* | | |
| | -0.0256* | 0.0027* | | 0.0007* | |
| | -0.0355* | 0.0030* | | | -0.0026* |
| | -0.0164* | 0.0026* | -0.0078* | 0.0006* | -0.0072* |
| Drift | | | | | |
| | -0.1261* | 0.0007 | -0.0083 | | |
| | -0.1133* | 0.0007 | | 0.0015* | |
| | -0.1339* | 0.0006 | | | -0.0117* |
| | -0.0348* | 0.0008 | -0.0061 | 0.0121* | -0.0146* |
| First half of the drift | | | | | |
| | -0.0701* | 0.0007 | -0.0069* | | |
| | -0.0668* | 0.0004 | | 0.0007* | |
| | -0.0794* | 0.0009 | | | -0.0090* |
| | -0.0084* | 0.0002 | -0.0055* | 0.0084* | -0.0106* |
| Second half of the drift | | | | | |
| | -0.0939* | 0.0001 | -0.0071* | | |
| | -0.0848* | 0.0007 | | 0.0011* | |
| | -0.1005* | 0.0006 | | | -0.0079* |
| | -0.0757* | 0.0005 | -0.0053 | 0.0090* | -0.0108* |

Notes: We estimate the model, and the cumulative abnormal return model, $|CAR_{kad}| = \alpha + \beta_p PI_{kad} + \beta_s |S_{kad}| + \beta_{spread} Spread_{kad} + \beta_{vol} Vol_{kad} + \beta_c C_{kad} + \varepsilon_{kad}$, over forty days after the announcement, $|CAR_{kad}|$. We report the coefficient values, and we mark with an asterisk those coefficients that are statistically significant at the five percent level, using bootstrap standard errors. We ran a weighted least squares regression to control for the heteroskedasticity in the data since $\mathrm{Var}(\varepsilon_{kd}) \propto \frac{1}{C_{kd}^2}$, where $C_{kd}$ = market capital

**Table 13**
Pooled Data Analysis: Top Three Decile

| | $\beta_p$ | $\beta_s$ | $\beta_{spread}$ | $\beta_{vol}$ | $\beta_c$ |
|---|---|---|---|---|---|
| Two-Day Abnormal Return | | | | | |
| | -0.0284* | 0.0041* | -0.0066* | | |
| | -0.0285* | 0.0038* | | 0.0005* | |
| | -0.0365* | 0.0043* | | | -0.0031* |
| | -0.0210* | 0.0036* | -0.0048* | 0.0053* | -0.0065* |
| Drift | | | | | |
| | -0.1565* | 0.0007 | -0.0128 | | |
| | -0.1542* | 0.0005 | | 0.0011* | |
| | -0.1735* | 0.0012 | | | -0.0139* |
| | -0.0605* | 0.0029 | -0.0085 | 0.0128* | -0.0160* |
| First half of the drift | | | | | |
| | -0.0549* | 0.0002 | -0.0086 | | |
| | -0.0587* | 0.0001 | | 0.0004* | |
| | -0.0687* | 0.0005 | | | -0.0096* |
| | -0.0480* | 0.0005 | -0.0060 | 0.0079* | -0.0103* |
| Second half of the drift | | | | | |
| | -0.1320* | 0.0004 | -0.0067 | | |
| | -0.1279* | 0.0003 | | 0.0008* | |
| | -0.1406* | 0.0003 | | | -0.0101* |
| | -0.0693* | 0.0007 | -0.0039 | 0.0083* | -0.0103* |

Notes: We repeat the same exercise as before only using observations in the top three surprise deciles. We estimate the two-day abnormal returns model, $|AR_{kd}| = \alpha + \beta_p PI_{kd} + \beta_s |S_{kd}| + \beta_{spread} Spread_{kd} + \beta_{vol} Vol_{kd} + \beta_c C_{kd} + \varepsilon_{kd}$, and the cumulative abnormal return model, $|CAR_{kd}| = \alpha + \beta_p PI_{kd} + \beta_s |S_{kd}| + \beta_{spread} Spread_{kd} + \beta_{vol} Vol_{kd} + \beta_c C_{kd} + \varepsilon_{kd}$, over forty days after the announcement, $|CAR_{kd}|$ , over the first twenty days after the earnings announcement, $|CAR_{1kd}|$ and over the last twenty days after the earnings announcement, $|CAR_{2kd}|$. We report the coefficient values, and we mark with an asterisk those coefficients that are statistically significant at the five percent level, using bootstrap standard errors. We ran a weighted least squares regression to control for the heteroskedasticity in the data since $\mathrm{Var}(\varepsilon_{kd}) \propto \frac{1}{C_{kd}^2}$, where $C_{kd}$ = market capital

**Table 14**

Pooled Data Analysis: Bottom Three Deciles

| | $\beta_p$ | $\beta_s$ | $\beta_{spread}$ | $\beta_{vol}$ | $\beta_c$ |
|---|---|---|---|---|---|
| Two-Day Abnormal Return | | | | | |
| | -0.0146* | 0.0017* | -0.0115* | | |
| | -0.0076 | 0.0011 | | 0.0012* | |
| | -0.0241* | 0.0016 | | | -0.0012* |
| | -0.0208* | 0.0010 | -0.0093* | 0.0079* | -0.0094* |
| Drift | -0.1206* | 0.0034 | -0.0085 | | |
| | -0.0789* | 0.0016 | | 0.0029* | |
| | -0.1139* | 0.0025 | | | -0.0061* |
| | -0.0737* | 0.0015 | -0.0036 | 0.0141* | -0.0154* |
| First half of the drift | | | | | |
| | -0.0736* | 0.0024* | -0.0069 | | |
| | -0.0463* | 0.0012 | | 0.0020* | |
| | -0.0720* | 0.0019 | | | -0.0051* |
| | -0.0416* | 0.0011 | -0.0033 | 0.0114* | -0.0129* |
| Second half of the drift | | | | | |
| | -0.0924* | 0.0020 | -0.0109* | | |
| | -0.0686* | 0.0008 | | 0.0021* | |
| | -0.0939* | 0.0014 | | | -0.0044* |
| | -0.0599* | 0.0007 | -0.0075 | 0.0102* | -0.0113* |

Notes: We repeat the same exercise as before only using observations in the bottom three surprise deciles. We estimate the two-day abnormal returns model, $|AR_{kd}| = \alpha + \beta_p PI_{kd} + \beta_s |S_{kd}| + \beta_{spread} Spread_{kd} + \beta_{vol} Vol_{kd} + \beta_c C_{kd} + \varepsilon_{kd}$, and the cumulative abnormal return model, $|CAR_{kd}| = \alpha + \beta_p PI_{kd} + \beta_s |S_{kd}| + \beta_{spread} Spread_{kd} + \beta_{vol} Vol_{kd} + \beta_c C_{kd} + \varepsilon_{kd}$, over forty days after the announcement, $|CAR_{kd}|$, over the first twenty days after the earnings announcement, $|CAR_{1kd}|$, and over the last twenty days after the earnings announcement, $|CAR_{2kd}|$. We report the coefficient values, and we mark with an asterisk those coefficients that are statistically significant at the five percent level, using bootstrap standard errors. We ran a weighted least squares regression to control for the heteroskedasticity in the data since $\mathrm{Var}(\varepsilon_{kd}) \propto \frac{1}{C_{kd}^2}$, where $C_{kd}$ = market capital

**Table 15**

Pooled Data Analysis: Past Returns Effects

| | $\beta_p$ | $\beta_s$ | $\beta_{spread}$ | $\beta_{vol}$ | $\beta_c$ | $\lvert AR_{kd}\rvert$ | $\lvert CAR_{kd-1}\rvert$ |
|---|---|---|---|---|---|---|---|
| Two-Day Abnormal Return | -0.0152* | 0.0026* | -0.0078* | 0.0006* | -0.0071* | | -0.004 |
| Drift | -0.1017* | 0.0007 | -0.0046 | 0.0012* | -0.0131* | 0.3144* | -0.012 |
| First half of the drift | -0.0549* | 0.0002 | -0.0047 | 0.0005* | -0.0094* | 0.2433* | -0.022* |
| Second half of the drift | -0.0753* | 0.0007 | -0.0049 | 0.0009* | -0.0102* | 0.1726* | -0.010 |

Notes: We estimate the model, $\lvert AR_{kd}\rvert = \alpha + \beta_p PI_{kd} + \beta_s \lvert S_{kd}\rvert + \beta_{spread} Spread_{kd} + \beta_{vol} Vol_{kd} + \beta_c \lvert C_{kd}\rvert + \beta_{CAR} \lvert CAR_{kd-1}\rvert + \varepsilon_{kd}$, and the cumulative abnormal return model, $\lvert CAR_{kd}\rvert = \alpha + \beta_p PI_{kd} + \beta_s \lvert S_{kd}\rvert + \beta_{spread} Spread_{kd} + \beta_{vol} Vol_{kd} + \beta_c \lvert C_{kd}\rvert + \beta_{AR} \lvert AR_{kd}\rvert + \beta_{CAR} \lvert CAR_{kd-1}\rvert + \varepsilon_{kd}$, over forty days after the announcement, $\lvert CAR_{kd}\rvert$. We report the coefficient values, and we mark with an asterisk those coefficients that are statistically significant at the five percent level, using bootstrap standard errors. We ran a weighted least squares regression to control for the heteroskedasticity in the data since $\mathrm{Var}(\varepsilon_{kd}) \propto \frac{1}{C_{kd}^2}$, where $C_{kd}$ = market capital

**Table 16**

Two-Day Abnormal Return and Post-Announcement Drift: Low Probability and High Probability of PIBT

| | Low Probability of PIBT | | High Probability of PIBT | |
|---|---|---|---|---|
| Decile | 2-Day Ab. Returns % | Drift % | 2-Day Ab. Returns % | Drift % |
| 1 | -0.94* | -1.73* | -0.22 | -0.65 |
| 2 | -0.80* | -0.89* | -0.17 | -0.43 |
| 3 | -0.36 | -0.45 | -0.15 | -0.11 |
| 4 | -0.23 | -0.44 | -0.03 | 0.11 |
| 5 | -0.09 | -0.42 | 0.16 | 0.20 |
| 6 | 0.01 | 0.06 | 0.67 | 0.35 |
| 7 | 0.44 | 0.63 | 0.53 | 0.41 |
| 8 | 1.00* | 0.77* | 0.36 | 0.73 |
| 9 | 1.13* | 0.84* | 0.62 | 0.74 |
| 10 | 1.55* | 1.79* | 0.73 | 0.95* |

Notes: We present in this table the means of variables, categorized by surprise deciles. The surprise decile is a ranking from 1 to 10 of this quarter's surprise $S_{kt}$ (firm k, quarter t), based on the previous quarter's surprise decile cut offs. The two-day abnormal return is the cumulative size adjusted return on the day of the announcement and the day after. We look at the day after, because some earnings announcements take place after the stock market is closed and the price reaction is observed the next day. Drift is calculated as the cumulated size adjusted return over trading days [+1,+40], where day 0 is the day after the earnings announcement date. The asterisk denotes that the abnormal returns are statistically significant at the five percent level. The high probability (low probability) of PIBT is defined as days when the probability of PIBT prior to the earnings announcement was above the 80$^{th}$ percentile (below the 20$^{th}$ percentile).