What moves stock prices?
The role of news, noise, and information☆

Jonathan Brogaard [a], Huong Nguyen [b], Talis J. Putnins [b, c], Eliza Wu [d]

[a] *University of Utah*
[b] *University of Technology Sydney*
[c] *Stockholm School of Economics in Riga*
[d] *University of Sydney Business School*

September 2019

Abstract

We develop a return variance decomposition model to separate the role of different types of information and noise in stock price movements. We disentangle four components: market-wide information, private firm-specific information revealed through trading, firm-specific information revealed through public sources, and noise. Overall, 31% of the return variance is from noise, 37% from public firm-specific information, 24% from private firm-specific information, and 8% from market-wide information. Since the mid-1990s, there has been a dramatic decline in noise and an increase in firm-specific information, consistent with increasing market efficiency.

JEL classification: G12; G14; G15

Keywords: variance decomposition; firm-specific information; market-wide information; stock return synchronicity

☆ We thank the Securities Industry Research Centre of Asia-Pacific (SIRCA) and Thomson Reuters for providing access to data used in this study. We also thank Paul Besson, Jean-Philippe Bouchaurd, Amit Goyal, Alok Kumar, Charles-Albert Lehalle, Joshua Mollner, Ryan Riordan, Elvira Sojli, and seminar and conference participants at the Market Microstructure Confronting Many Viewpoints #5 Conference, the Investment Management Research Conference, the Asian Finance Association Annual Meeting, the Financial Research Network (FIRN) Annual Meeting; FIRN Asset Pricing Workshop, the Australasian Finance and Banking Conference, the J.P. Morgan Quantference, Brigham Young University, Monash University, the University of Western Australia, Curtin University, the University of Technology Sydney, and the University of Sydney for comments that have helped to improve this paper.

The issue of what drives stock price movements is a fundamental question in finance with implications for understanding risk, informational efficiency, and asset pricing. By understanding the stock return generating process, researchers can address questions such as whether private information is more important than public information, whether the role of market-wide information is increasing or decreasing over time, or how much noise is in stock price movements. A methodology for measuring the return generating process is also useful for evaluating the impacts of recent phenomena such as the growth in passive investing and algorithmic trading, among others. This paper develops a new tool that allows stock returns to be decomposed into various information components while simultaneously allowing for price changes to occur due to non-informational reasons.

There are currently two dominant approaches to decomposing the drivers of stock price changes. One approach exploits the canonical discounted cash flow valuation model to divide a return series into cash flow and discount rate related return components (Campbell and Shiller, 1988a, 1988b; Campbell, 1991). The second decomposes returns into market-wide news and firm-specific news using the R^2 from a regression of stock returns on market returns (Morck, Yeung, and Yu, 2000).

The existing decompositions have limitations, which we overcome with the variance decomposition developed in this paper. To illustrate, Figure 1 shows the time-series of the R^2 measure. The R^2 time-series suggests that market efficiency has been on the decline since the mid-1990s (higher R^2 implies less firm-specific information is reflected in prices). This is at odds with much of the event study literature that suggests different market changes have generally improved market efficiency in recent years (e.g., Brogaard, Hendershott, and Riordan, 2014; Comerton-Forde and Putnins, 2015).

Insert Figure 1 About Here

The first limitation of existing decompositions arises from ignoring the role of noise in stock returns. Temporary deviations of prices from their equilibrium levels are part of the return generating process (e.g., Hendershott and Menkveld, 2014; Asparouhova, Bessembinder, and Kalcheva, 2010; Asparouhova, Bessembinder, and Kalcheva, 2013). These price deviations, or "noise," arise from microstructure frictions such as bid-ask spreads, nonsynchronous trading, discrete price grids, and temporary price impacts of order imbalances, as well as changes in investor sentiment or other behavioral factors, in combination with limits to arbitrage (Asparouhova et al., 2013). Noise has a significant effect on returns at daily and monthly frequencies, not just intraday horizons. For example, Blume and Stambaugh (1983) and Asparouhova et al. (2013) show that noise at daily frequencies causes an economically meaningful bias in returns, equal to 50% or more of the corrected estimate and is able to explain the size effect. Jegadeesh (1990) and Lehmann (1990) document significant reversals in stock returns at monthly and weekly horizons, respectively, also consistent with the notion that daily, weekly, and monthly returns contain substantial noise.

Noise can distort the existing measures. For instance, an increase in R^2 is generally interpreted as indicating a relative decrease in the amount of firm-specific information in prices. However, if returns also contain noise, an alternative interpretation of a high R^2 is that there has been a decrease in the magnitude of idiosyncratic pricing errors and therefore an increase in R^2 does not necessarily indicate deteriorating informational efficiency.

The second limitation is the inability of existing methods to disentangle information into more refined categories. The partition into market-wide and firm-specific information has long been of interest in finance as firm-specific information is vital for efficient resource allocation across firms. Some recent studies suggest the rapid growth in passive investing could harm the firm-specific information in stock prices (e.g., Cong and Xu, 2017). Similar concerns have been raised about the effects of high-frequency trading (e.g., Baldauf and Mollner, 2018). Further, the partition into private and public information is important given substantial changes in the regulation of corporate information disclosure in recent decades (e.g., Regulation Fair Disclosure (2000) and the Sarbanes Oxley Act (2002)). Such regulations could result in better disclosure crowding out private information acquisition, with implications for the profitability of active investing and the levels of adverse selection. Having a variance decomposition methodology to shed light on cross-sectional variation and general time trends can aid research on these types of information topics.

We propose a new return variance decomposition model that explicitly accounts for noise and partitions information into various sources. For example, in the baseline model, we distinguish between market-wide information, firm-specific information revealed through trading on private information, and firm-specific information revealed through public news. In an extended version of the model, we further decompose the information components into cash flow and discount rate sub-components to relate the model to this long-standing branch of the asset pricing literature. Our approach allows for a more nuanced understanding of the specific sources of information that is impounded into stock prices along economically meaningful dimensions. It also allows the variance decompositions to be performed at higher frequencies (e.g., annual decompositions of daily returns) and therefore allows researchers to examine time-series variation in the components of stock return variance.[1]

We motivate our approach with an extension of the model of Jin and Myers (2006), allowing for noise traders and pricing errors. We allow returns to be driven by noise, firm-specific information revealed by trading on private information or by other sources such as public news, and market-wide information.

Our empirical return variance decomposition model draws on the market microstructure toolkit where separating temporary price movements driven by frictions from permanent price movements is commonplace. For example, Hasbrouck (1993) separates noise from information through a temporary-permanent component decomposition. Permanent innovations are those that affect the long-run expected value of the security, whereas temporary innovations affect prices in the short-run, but have no effect on the long-run expected value of the security. Permanent innovations in prices reflect innovations in the fundamental value driven by new information, whereas temporary innovations are pricing errors. These deviations from the fundamental value can arise from various well-documented market frictions including the bid-ask bounce, discrete price grids, temporary price pressures created by uninformed buying or selling, and non-synchronous trading.

[1] Most existing variance decompositions rely on low frequencies to mitigate the effects of noise, precluding time-series analysis.

We build on Hasbrouck's original decomposition by adapting the model to daily returns and separating innovations in the fundamental value into market-wide information, firm-specific information revealed through trading on private information, and firm-specific information revealed through other sources (public information). This gives rise to four components of return variance that map to the theoretical model. We estimate the model using daily returns on all common stocks listed on the NYSE, AMEX, and NASDAQ between 1960 and 2015, performing the variance decomposition separately for every stock in every year. This approach minimizes the issue of non-stationarity and allows us to examine how the variance components change in the cross-section and through time.

Intriguingly, we find that roughly 31% of daily return variance is noise. Firm-specific information accounts for the majority (61%) of stock return variance, with market-wide information accounting for the remaining 8% of variance in the full sample. We further partition firm-specific information and find that in the full sample, public firm-specific information plays a larger role than private firm-specific information that is impounded into prices through trading (37% and 24% of variance, respectively). While the estimates suggest that noise makes up an economically meaningful share of daily stock return variance, the estimate is substantially lower than estimates of noise at intra-day horizons (82%).[2] Our estimate is consistent with Hendershott and Menkveld (2014) who find that the ratio of "price pressure," distortions of the midquote price from the efficient price, to the variance of efficient prices is 33% in their sample of 697 NYSE stocks during 1994-2005. It is also consistent with Hendershott et al. (2011) who estimate that one-quarter of monthly return variance in NYSE stocks is due to transitory price changes explained by order imbalances and market-makers' inventories. Finally, the estimate is also consistent with the economically meaningful return reversals at daily, weekly, and monthly horizons, which reflect temporary departures from efficient prices due to imperfect liquidity (e.g., Jegadeesh, 1990; Lehmann, 1990; Avramov, Chordia, and Goyal, 2006; Nagel, 2012).

[2] Extrapolating from Hasbrouck (1991b, 1993) for the year 1989.

We find substantial time-series variation in the components of variance. Some key trends stand out. First, noise increases from the 1970s to the mid-1990s, in particular around a period of collusion by dealers that widened bid-ask spreads. The subsequent decline in noise corresponds to a period with a general improvement in liquidity and exogenous decreases in tick sizes (minimum price increments). Separating the sample by firm size shows that the noise component decreases monotonically with firm size. Larger firms have less noise in their prices, as expected. When separating the sample by industry we observe only minor fluctuations in the different components of variance, suggesting that the findings are not specific to a particular industry, nor are they driven by a certain segment of the economy.

Second, the role of firm-specific information has increased through time, driven largely by increases in the amount of public firm-specific information that is reflected in prices. This trend is consistent with increasing informational efficiency through time, which one may expect given a variety of regulations such as the Regulation Fair Disclosure (2000) and the Sarbanes Oxley Act (2002) that have increased the quality and quantity of corporate disclosure. Third, market-wide information has become a less important driver of stock returns over time. While market-wide information tends to spike during crises, it has generally declined from around 15% of variance to around 5-10% in recent years.

During the 55-year sample period markets have changed dramatically. The number of exchanges has increased, bid-ask spreads have decreased, intermediaries have transitioned from dealers to market-makers to high-frequency traders, and investor horizons have declined. In addition, the quantity of information has ballooned and at the same time the cost and latency of accessing it has plummeted. These time-varying changes can impact the various components of the price process differently. Because we re-calculate the decomposition each year for each stock, our measure is able to vary along with these changes in the market.

We provide three detailed examples of how this decomposition can be useful for examining prominent questions in finance. First, we look at the role tick size has on the noise component. The tick size was reduced from eighths of a dollar to sixteenths of a dollar in 1997 and then from sixteenths to pennies in 2001. Chordia, Roll, and Subrahmanyam (2008) show that informational efficiency improved significantly around the change from quoting in eighths of a dollar to sixteenths, with the second event having a smaller effect. We exploit this natural experiment and examine how the noise share of variance is affected by the reduction in tick size from eighths of a dollar to sixteenths. We find that following the tick size reduction there is a substantial decrease in the noise in stock returns. Stocks experiencing a larger change in their relative tick size exhibit a larger decrease in noise.

Second, we examine exogenous shocks to analyst coverage. We find that a decline in analyst coverage harms the quality of the corporate information environment and reduces the proportion of public firm-specific information in prices. We use brokerage mergers and closures as plausibly exogenous shocks to analyst coverage because the termination of coverage is not driven by the characteristics or behavior of the firm (see Hong and Kacperczyk, 2010; Kelly and Ljungqvist, 2012; Brogaard et al., 2018). Additionally, exogenous decreases in analyst coverage are also associated with an increase in the relative level of noise in prices.

Third, we compare our variance components to two other measures of information in prices: the Hou and Moskowitz (2005) delay metric and the variance ratio (French and Roll, 1986; Chordia, Roll, and Subrahmanyam, 2011). The first measure focuses specifically on market-wide information, while the second measure focuses on information generally. We find a strong inverse relation between the delay metric and the share of variance that is attributable to market-wide information in our variance decomposition, whereas we find the variance ratio is positively related to all of the information components of variance and negatively related to the amount of noise in prices. The results of the decomposition reassuringly fit with what the existing, less flexible measures would predict.

The final set of analysis we conduct reconciles our decomposition methodology with the large return decomposition literature on the relative importance of cash flow versus discount rate news (e.g., Campbell and Ammer, 1993; Vuolteenaho, 2002; Chen et al., 2013). After accounting for noise in the return generating process, we find a considerably smaller role for discount rate news and a much larger role for cash flow news. Without accounting for noise we find the ratio of cash flow news to discount rate news for individual stocks is around five, similar to Vuolteenaho (2002). After accounting for noise the ratio increases to around 25 because at the individual stock level the discount rate news accounts for only a small (3%) fraction of the stock return variance.

The difference in the importance of discount rate news compared to previous studies stems largely from the fact that noise induces return predictability. For example, returns have negative first-order serial correlation at many frequencies due to pricing errors and price pressure (e.g., Jegadeesh, 1990; Roll, 1984). Without accounting for noise, the noise-induced variation in predicted returns is attributed to variation in the discount rate thereby overstating the role of discount rate news. We show the ratio of cash flow news to discount rate news is higher in firm-specific information than in market-wide information. Therefore, although cash flow news is generally responsible for a larger share of stock return variation than discount rate news, cash flow news tends to be more idiosyncratic and discount rate news more systematic. This result helps reconcile existing studies in which cash flow news is found to be more important in stock-level analyses (e.g., Vuolteenaho, 2002), while discount rate news plays a larger role in portfolio-level analyses in which much of the idiosyncratic variation is removed (e.g., Campbell, 1991).

Being able to accurately measure the amount and type of information in asset prices is vital to understanding the impacts of recent trends and innovations in finance. This paper makes a methodological contribution to the literature by developing a richer and more general variance decomposition that allows the separation of variance into multiple information and noise components.

1. Related literature

This paper relates to three main bodies of literature. The first is stock return variance decompositions. For example, Campbell and Shiller (1988a, 1988b) and Campbell (1991) decompose the variance of unexpected stock returns into two components: cash flow news and discount rate news. Chen, Da, and Zhao (2013) also decompose stock return variance into these components but with a method that directly incorporates cash flow forecasts. A different partition is used in Campbell et al. (2001), who decompose stock return variations into three components: a market-wide return, an industry-specific residual, and a firm-specific residual. A simple but widely used method for stock return decomposition is offered by Roll

(1988), who simply distinguishes market-wide variations and firm-specific variations. Our paper differs from these existing methods by using random walk variance decompositions, drawn from the market microstructure literature, to separate information from noise, and partition information into a more granular set of components. This separation of noise and information provides a measure related to market efficiency and decomposes the contributions of different information sources.

Second, our study is related to a growing body of literature on the significant impact of noise on asset prices and returns. For instance, Blume and Stambaugh (1983) show that zero-mean noise in prices leads to a positive bias in mean returns. Asparouhova et al. (2010) find that noisy prices lead to biases in intercept and slope coefficients obtained in any OLS regression using return as the dependent variable. In addition, Asparouhova et al. (2013) find that correcting for the effects of noise in prices has significant effects on return premium estimates from monthly return data. Motivated by this strand of literature, we incorporate noise into our model as an important factor that might have impacts on stock price movements. In turn, by providing a convenient method to estimate noise in prices, our approach provides a tool that enables future research to more systematically examine the drivers and effects of noise.

Third, our paper contributes to the extensive literature measuring market efficiency. For instance Bai, Philippon, and Savov (2016) use quarterly cross-sectional regressions of the extent to which market prices predict earnings as a measure of market-wide efficiency. In contrast, our method provides a more granular measure at the individual *stock* level, rather than at the market level, and does not require earnings information that may be subject to reporting bias.

There are several traditional efficiency measures that are largely based on the concept of weak-form market efficiency, including autocorrelations, variance ratios, reversal strategies, delay measures, post-earnings drift, profitability of momentum strategies, and intraday return predictability based on past order flow or past returns. Rosch, Subrahmanyam, and van Dijk (2017) examine the dynamics of market efficiency in the United States using the first principal component of four existing intra-day efficiency measures, while Griffin, Kelly, and Nardari (2010) analyze several of these traditional measures across international markets. Interestingly, Griffin et al. (2010) find that momentum trading strategies are more profitable in developed markets, prices deviate more from a random walk in developed markets, and prices in emerging markets incorporate past market returns more quickly than prices in developed markets. Thus, the existing measures produce results that are inconsistent with the conventional wisdom that emerging markets are less efficient than developed markets. Griffin et al. (2010) argue that their results highlight crucial limitations of traditional weak-form efficiency measures and point to the importance of measuring informational aspects of efficiency.

Our approach follows this call by focusing on decomposing the *information* in prices rather than contributing to the set of weak-form efficiency measures. Besides being based on the broader asset pricing methodology, our measure also differs from existing approaches, including Bai, Philippon, and Savov (2016), in that it not only measures efficiency in terms of information versus noise, but also decomposes the specific sources of information that is reflected in prices. The decomposition of information provides a more complete picture of the nature of market efficiency and how it evolves through time and in the cross-section.

2. Empirical model for variance decomposition

This section lays out the empirical model that we use to separate noise and various sources of information. It begins with a theoretical motivation for the components of variance that we then empirical separate in the data.

2.1. Model motivation

To understand the different sources of variation in stock prices, we derive a modified version of the Jin and Myers (2006) model. In their original model, the intrinsic value of the firm is the present value of future operating cash flows. Cash flows are affected by market-wide and firm-specific shocks. Market participants have perfect information about these shocks, other than the private firm-specific shocks, which are not disclosed to the market. In this setup, Jin and Myers (2006) show that the R^2 in a regression of a firm's returns on those of the market is decreasing in the transparency of the firm because when more of the private firm-specific shocks are disclosed to the market, the stock-specific information in returns increases, thereby reducing the R^2.

We extend this setup by adding noise to stock returns. We inject noise into the model as one of the shocks to cash flow information. This has several interpretations. First, it could be viewed as taking the perfect information that is provided to the market in the original model and making it imperfect such that estimation error of market participants induces noise in their perceptions of fundamental information. It could also be interpreted as the addition of noise traders to the model in the spirit of Black's (1986) notion that "noise trading is trading on noise as if it were information." Finally, the ultimate result of injecting noise is that returns vary around the efficient returns, which in real markets can occur due to frictions such as a discrete pricing grid, non-synchronous trading, or imperfect liquidity. A second departure from Jin and Myers (2006) is to split firm-specific information into a part that is revealed through trading on private information, and a part revealed through public information such as company announcements and news.

What these extensions to the Jin and Myers (2006) model show is that stock return variance can be decomposed into four distinct sources of variation, which can be measured by four variance shares:

$$\eta_j = \frac{Var(\varepsilon_{j,t})}{Var(\varepsilon_{1,t} + \varepsilon_{2,t} + \varepsilon_{3,t} + \varepsilon_{4,t})}, \tag{1}$$

with $j = \{1, 2, 3, 4\}$. The first of these variance shares, η_1, is the contribution of market-wide information to an individual stock's variance. The second (η_2) is the contribution of firm-specific information that is revealed through trading (private firm-specific information), while the third (η_3) is the contribution of public firm-specific information. The last component (η_4) is the effect of noise on stock return variance. Our empirical model seeks to estimate these four variance shares. The extended version of this model and the proofs are in Appendix A.

2.2. *Baseline variance decomposition model*

To empirically estimate the variance shares from Equation (1), we propose a variance decomposition model that separates noise and various sources of information. Our approach to separating noise from information builds on Hasbrouck (1993), who shows how temporary-permanent decompositions can be used on stock returns to separate temporary pricing errors from innovations in the fundamental value.[3] Permanent innovations in stock prices are those that affect the long-run expected value of the security, whereas temporary innovations affect prices in the short-run, but have no effect on the long-run expected value of the security. Permanent innovations in prices therefore reflect innovations in the fundamental value driven by new information being impounded into prices, whereas temporary innovations are pricing errors (deviations from fundamentals) generically referred to as "noise." Noise is caused by many factors including the bid-ask spread, discrete price grids, illiquidity, temporary price pressures created by uninformed buying or selling, and non-synchronous trading.

Our variance decomposition departs from Hasbrouck (1993) in two important ways. First, we adapt the approach so that it can be applied at lower frequencies such as daily returns (Hasbrouck (1993) models intraday trade-to-trade returns). We undertake an array of validation tests to verify that the variance decomposition produces reasonable estimates. Second, we push the variance decomposition further to separate the information into market-wide information, firm-specific information revealed through trading on private information, and firm-specific information revealed through public information. This gives rise to four components of variance that map to the theoretical model. These components are illustrated in Figure 2.

Insert Figure 2 About Here

Consider p_t, the logarithm of the observed price at time t, as the sum of two components:

[3] Similar temporary-permanent decompositions are also used in empirical macroeconomics.

$$p_t = m_t + s_t \ , \tag{2}$$

where m_t is the efficient price and s_t is the pricing error. The pricing errors can have a temporary (short-run) affect on the price, but they do not affect price in the long run (no permanent effect). m_t follows a random walk with drift μ, and innovations w_t:

$$m_t = m_{t-1} + \mu + w_t \ . \tag{3}$$

The innovations reflect new information about the stock's fundamentals and are thus unpredictable, $\mathrm{E}_{t-1}[w_t] = 0$. The drift is the discount rate on the stock over the next period (day).[4] The stock return is therefore:

$$r_t = p_t - p_{t-1} = \mu + w_t + \Delta s_t \ . \tag{4}$$

Suppose that there are three sources of information impounded into stock prices: market-wide information, private firm-specific information incorporated through trading, and public firm-specific information such as firm-specific news disseminated in company announcements and by the media. The random-walk innovations, w_t, in Equation (3) can then be decomposed into three parts:

$$w_t = \beta \varepsilon_{r_m,t} + \delta \varepsilon_{x,t} + u_t \ , \tag{5}$$

and thus

$$r_t = \mu + \beta \varepsilon_{r_m,t} + \delta \varepsilon_{x,t} + u_t + \Delta s_t \ , \tag{6}$$

where $\varepsilon_{r_m,t}$ is the unexpected innovation in the market return, $\beta \varepsilon_{r_m,t}$ reflects the market-wide information incorporated into stock prices, $\varepsilon_{x,t}$ is an unexpected innovation in signed dollar volume, $\delta \varepsilon_{x,t}$ is the firm-

[4] Later, when separating cash flow and discount rate news, we allow the discount rate to be time-varying.

specific information revealed through trading on private information, and u_t is the remaining part of firm-specific information that is not captured by trading on private information. The separation of firm-specific information into private information associated with trading and public information not associated with trading follows Hasbrouck (1991a, 1991b). The pricing error, Δs_t, can be correlated with the innovations in the efficient price, w_t.

Equation (6) can be modeled as a vector auto-regression (VAR) to account for serial correlations in returns and other explanatory variables. We use a structural VAR with five lags to allow a full week of lagged effects:[5]

$$\begin{aligned} r_{m,t} &= a_0 + \textstyle\sum_{l=1}^{5} a_{1,l} r_{m,t-l} + \sum_{l=1}^{5} a_{2,l} x_{t-l} + \sum_{l=1}^{5} a_{3,l} r_{t-l} + \varepsilon_{r_m,t} \\ x_t &= b_0 + \textstyle\sum_{l=0}^{5} b_{1,l} r_{m,t-l} + \sum_{l=1}^{5} b_{2,l} x_{t-l} + \sum_{l=1}^{5} b_{3,l} r_{t-l} + \varepsilon_{x,t} \\ r_t &= c_0 + \textstyle\sum_{l=0}^{5} c_{1,l} r_{m,t-l} + \sum_{l=0}^{5} c_{2,l} x_{t-l} + \sum_{l=1}^{5} c_{3,l} r_{t-l} + \varepsilon_{r,t}, \end{aligned} \tag{7}$$

where $r_{m,t}$ is the market return, x_t is the signed dollar volume of trading in the given stock with positive values capturing net buying and negative values capturing net selling, and r_t is the stock return.[6] The lags of stock returns account for short-term momentum as well as reversals that can be driven by temporary price impacts from trading (e.g., Hendershott and Menkveld, 2014). The lags of signed dollar volume

[5] There are two approaches to resolving the contemporaneous relations between the variables in the VAR. One is to use a structural VAR that explicitly defines the assumptions about contemporaneous causality using economic arguments. The second is to use a reduced form VAR and apply Cholesky factorization to the variance-covariance matrix of reduced form innovations, which itself implies a recursive causal chain from the first to the last variable in the system. We take the former approach. The structural VAR approach has at least two advantages: (i) while both approaches must make assumptions about excluded contemporaneous relations to facilitate identification, the structural model allows the assumptions to be guided by economic reasoning rather than arbitrary allocation, and (ii) the structural approach produces a unique variance decomposition whereas the reduced form approach produces a decomposition that is sensitive to the ordering of variables in the model.

[6] Our measure of the market returns is the daily value-weighted market return excluding American Depository Receipts. We use market returns rather than a broader collection of factors such as market/size/value because we seek to identify the contribution of *market-wide* information to stock return variance. The variance decomposition could easily be extended to quantify the role of a variety of factor information in prices. We construct a proxy for the daily signed dollar volume of each stock as a product of price, volume, and sign of the return, similar to Pastor and Stambaugh (2003).

account for persistence in order flow (e.g., Hasbrouck 1988). Finally, the lags of market returns account for first-order serial correlation in market returns due to non-synchronous trading (e.g., Scholes and Williams, 1977) as well as delayed stock price reactions to market-wide information (e.g., Hou and Moskowitz, 2005).

The structural VAR above embeds contemporaneous relations between the variables. First, market-wide information can be reflected in stocks contemporaneously, but because each stock is a small part of the market index, individual stock returns and trades have a negligible contemporaneous impact on the market return. Second, trading activity in a stock can be contemporaneously caused by market returns and can contemporaneously cause changes in the stock price, but not vice versa. To the extent that returns can trigger trading activity contemporaneously (within the same day) or exogenous events such as company announcements can trigger both (trading and returns), the model will tend to overstate the extent to which trading activity drives returns. Therefore, in splitting firm-specific information into public and private components, our structural assumptions will tend to estimate the upper bound on private information and lower bound on public information. As a result of explicitly modelling the contemporaneous relations between variables, the structural VAR errors $\{\varepsilon_{r_m,t}, \varepsilon_{x,t}, \varepsilon_{r,t}\}$ are contemporaneously uncorrelated.

We separately estimate the VAR for every stock-year using daily data. Keeping the estimation windows to one-year periods mitigates concerns about non-stationarity and allows the relations between variables to change through time, from one year to the next. Annual estimation also allows us to examine how the variance components change through time.

Next we transform the VAR in Equation (7) to an infinite order structural vector moving average (VMA) model:

$$
\begin{aligned}
r_{m,t} &= A_0 + \textstyle\sum_{l=0}^{\infty} A_{1,l}\varepsilon_{r_m,t-l} + \sum_{l=1}^{\infty} A_{2,l}\varepsilon_{x,t-l} + \sum_{l=1}^{\infty} A_{3,l}\,\varepsilon_{r,t-l} \\
x_t &= B_0 + \textstyle\sum_{l=0}^{\infty} B_{1,l}\varepsilon_{r_m,t-l} + \sum_{l=0}^{\infty} B_{2,l}\varepsilon_{x,t-l} + \sum_{l=1}^{\infty} B_{3,l}\,\varepsilon_{r,t-l} \\
r_t &= C_0 + \textstyle\sum_{l=0}^{\infty} C_{1,l}\varepsilon_{r_m,t-l} + \sum_{l=0}^{\infty} C_{2,l}\varepsilon_{x,t-l} + \sum_{l=0}^{\infty} C_{3,l}\,\varepsilon_{r,t-l}.
\end{aligned}
\tag{8}
$$

The VMA form of the model is useful to glean the intuition for the temporary-permanent decomposition that separates noise from the innovations in the efficient price. The permanent effect on a stock's returns from a shock to the arrival of market-wide information, $r_{m,t}$ (unanticipated market returns), is given by $\theta_{r_m} = \sum_{l=0}^{\infty} C_{1,l}$. This is also the cumulative impulse response tracing time forward to the point where the response to the shock stabilizes. Similarly, the permanent effect on a stock's returns from trading activity is $\theta_x = \sum_{l=0}^{\infty} C_{2,l}$, and the permanent effect from a shock to the stock's returns that is neither due to market-wide information nor trading is $\theta_r = \sum_{l=0}^{\infty} C_{3,l}$. We estimate θ_{r_m}, θ_x, and θ_r from the impulse response functions of the structural model.

The information-driven innovation in the efficient price is given by $w_t = \theta_{r_m}\varepsilon_{r_m,t} + \theta_x\varepsilon_{x,t} + \theta_r\varepsilon_{r,t}$. The efficient price drift is given by $\mu = C_0$. The innovation in the pricing error is given by $\Delta s_t = r_t - \mu - w_t = r_t - C_0 - \theta_{r_m}\varepsilon_{r_m,t} - \theta_x\varepsilon_{x,t} - \theta_r\varepsilon_{r,t}$. It follows that the variance of information-driven innovations in the efficient price is $\sigma_w^2 = \theta_{r_m}^2\sigma_{\varepsilon_{r_m}}^2 + \theta_x^2\sigma_{\varepsilon_x}^2 + \theta_r^2\sigma_{\varepsilon_r}^2$. Recall, the structural model errors are contemporaneously uncorrelated by construction and therefore the covariance terms are all zero. The contribution to the efficient price variation from each of the information components is $\theta_{r_m}^2\sigma_{\varepsilon_{r_m}}^2$ (market-wide information), $\theta_x^2\sigma_{\varepsilon_x}^2$ (private firm-specific information), and $\theta_r^2\sigma_{\varepsilon_r}^2$ (public firm-specific information). The variance of noise, σ_s^2, is computed from the time-series of Δs_t.

To examine the contribution of each component in the total stock return variance, we construct two groups of new measures, the first being contributions to variance:

$$\begin{aligned} MktInfo &= \theta_{r_m}^2\sigma_{\varepsilon_{r_m}}^2 \\ PrivateInfo &= \theta_x^2\sigma_{\varepsilon_x}^2 \\ PublicInfo &= \theta_r^2\sigma_{\varepsilon_r}^2 \\ Noise &= \sigma_s^2, \end{aligned} \tag{9}$$

and the second being shares of variance:

$$
\begin{aligned}
MktInfoShare &= \theta_{r_m}^2 \sigma_{\varepsilon_{r_m}}^2 / (\sigma_w^2 + \sigma_s^2) \\
PrivateInfoShare &= \theta_x^2 \sigma_{\varepsilon_x}^2 / (\sigma_w^2 + \sigma_s^2) \\
PublicInfoShare &= \theta_r^2 \sigma_{\varepsilon_r}^2 / (\sigma_w^2 + \sigma_s^2) \\
NoiseShare &= \sigma_s^2 / (\sigma_w^2 + \sigma_s^2).
\end{aligned}
\tag{10}
$$

Accordingly, $MktInfo$, $PrivateInfo$, $PublicInfo$, and $Noise$ are the variance contributions of market-wide information, trading on private firm-specific information, firm-specific information other than that revealed through trading, and noise, respectively. $PrivateInfoShare$, $PublicInfoShare$, and $MktInfoShare$ are corresponding shares of variance from those various sources of stock price movements.[7] These three shares sum to the contribution of the efficient price innovations (recall the covariances between these components are all zero by construction). Meanwhile, $NoiseShare$ reflects the relative importance of pricing errors due to illiquidity, price pressures, or other microstructure frictions.

As each stock may have different dynamics we perform the variance decomposition on each stock separately. In addition, because we examine a long time-series in which dynamics and the drivers of return variance might change through time, we conduct the analysis separately year by year. Therefore, we estimate the variance decomposition detailed above separately for each stock-year.

3. Variance components through time and in the cross-section

This section applies the variance decomposition model to the data. First, we describe the data (Section 3.1) and report the coefficients of the reduced form VAR (Section 3.2), which is one of the first steps in

[7] Rather than using the total variance of returns in the denominator of the variance shares, we instead take the sum of the efficient price innovations variance and the noise variance, which excludes the covariance between the efficient price component and Δs. In Appendix C we show that ignoring this covariance term has a negligible effect on the overall estimates of variance shares.

performing the variance decomposition. Next, we report the estimated variance components in the full sample and compare the noise in daily returns to that of intraday returns (Section 3.3). We then characterize how the variance components change through time (Section 3.4) and how they vary in the cross-section of stocks (Section 3.5).

3.1. Data

Our sample consists of all common stocks listed on the NYSE, AMEX, and NASDAQ. We use daily data on returns, prices, market capitalizations, volumes, and sectors for the period from 1960 to 2015 from the Center for Research in Security Prices (CRSP). Appendix B contains a summary of variable definitions.

3.2. VAR coefficients

Before proceeding to the variance decomposition, Table 1 reports reduced-form VAR coefficient estimates averaged across the individual VAR models.[8] Below each average, in parentheses, the Table reports the percentage of negative statistically significant (at 5%) coefficients (first number in the parentheses) and the percentage of positive statistically significant (at 5%) coefficients (second number in the parentheses). The coefficients in the market return equation (Panel A) show a tendency for positive first-order serial correlation in market returns, consistent with the effects of non-synchronous trading (Scholes and Williams, 1977) and slow diffusion of market-wide information. In contrast, lags of other variables (trading in individual securities and individual stock returns) do not explain current market returns.

The coefficients in the signed dollar volume equation (Panel B) show a tendency for buying to follow positive market returns. They also show positive serial correlation in daily signed dollar volume, consistent with persistence in order flow (e.g., Hasbrouck (1988) and many subsequent studies).

[8] The variance decomposition is based on the structural VAR in Equation (7), but for practical reasons we follow the common practice of first estimating the reduced-form VAR and then using the coefficients and reduced form error covariances to compute the structural VAR impulse response functions and structural VMA. These structural VAR impulse response functions and structural VMA are then used for the variance decomposition.

The coefficients in the individual stock return equation (Panel C) show that stock returns tend to be positively related to lagged market returns consistent with the known slow diffusion of market-wide information (e.g., Hou and Moskowitz, 2005). They also show a tendency in some stock-years for trading (innovations in signed dollar volume) to impact returns not only contemporaneously but also with a lag, suggesting that at times the information in trading takes more than one day to be fully reflected in prices. They also indicate the presence of negative serial correlation in stock returns at daily frequencies out to approximately four days, consistent with reversals of pricing errors due to price pressure (e.g., Hendershott and Menkveld, 2014). The VAR coefficient estimates support the use of five lags in the VAR because by the fifth lag very few coefficients are statistically different from zero (besides the 5% that would be expected by chance at the 95% confidence level).

In addition to the average coefficients, which reveal the lead-lag relations between variables, Table 1 also reports the average correlations of the reduced-form VAR residuals for pairs of variables. These correlations reveal the contemporaneous relations between innovations in the variables. Innovations in signed dollar volume are contemporaneously correlated with returns of individual stocks and with market returns. These correlations are consistent with buying pressure pushing prices up as well as positive returns inducing buying (and vice versa for negative returns). There is also a positive contemporaneous correlation between individual stock returns and market returns consistent with individual stocks contributing to the market return but also market returns reflecting market-wide information, which is impounded in individual stock prices.

Insert Table 1 About Here

3.3. Full sample estimates of variance components

Table 2 Panel A reports the estimated variance shares from the baseline model above for the full sample (all US stocks from 1960 to 2015). Recall the variance decomposition is performed separately for every stock-year. From the stock-year estimates we calculate variance-weighted averages of each component.[9] The results show that market-wide information is the smallest component and accounts for around 8% of stock return variance, while firm-specific information accounts for 61% (summing $PrivateInfoShare$ and $PublicInfoShare$). Most of the firm-specific information is impounded in prices through public information (37% of variance), while firm-specific private information that is impounded through trading accounts for around 24% of variance. Finally, noise accounts for a fairly substantial 31% of overall daily stock return variance.

Insert Table 2 About Here

Before exploring the time-series and cross-sectional patterns in these variance shares, we consider how these estimates, in particular the noise in returns, compare to other estimates. The first comparison is with intraday returns (e.g., trade-to-trade), which is where similar temporary-permanent decompositions were first used to separate noise from information. Extrapolating from Hasbrouck (1991b, 1993), the implied noise share in intraday trade-to-trade returns is around 82% in US stocks in 1989.[10] For direct comparison, in the year 1989, our model estimates that the noise share in daily returns is around 35%. Therefore, the estimates of the level of noise in daily returns are considerably smaller than estimates of the noise in intraday trade-to-trade returns. One of the reasons for why there is less noise in daily returns than in trade-to-trade returns is that some sources of noise, such as the bid-ask spread, do not scale up when the

[9] We use variance-weighted averages in our baseline results for comparability with Morck et al. (2000, 2013). The motivation is that we are trying to understand the drivers of return variance and so want to emphasize stocks that experience greater variance. Alternatively, using equal-weighting across stocks produces similar results (trends and levels of the components). Using value-weighting across stocks produces lower estimates of the level of noise (16% lower in the pooled sample) and higher estimates of the information components of variance, consistent with the notion that smaller stocks have noisier prices. More direct evidence on the cross-sectional differences in the variance components is in Section 3.4.

[10] Hasbrouck (1991, 1993) does not report a noise share comparable to ours but we are able to calculate one from his results as follows. The estimated variance of pricing errors in Hasbrouck (1993) is 10.89×10^{-6}, whereas the variance of random walk innovations in Hasbrouck (1991b) using the same sample is 4.7×10^{-6}. If we conservatively assume zero serial correlation in pricing errors (such an assumption will underestimate the noise variance in the following calculation), then the variance of changes in pricing errors is $2\times10.89\times10^{-6}$ (it is the variance of *changes* in pricing errors that adds noise to returns, not just variance of pricing errors). Now if we compute an implied noise share we get 82.17%.

return horizon is increased, yet fundamental volatility (the variance attributable to information) does scale up with the return horizon (this intuition is exploited in the Corwin and Schultz (2012) effective spread estimator). For example, a one-minute return between two successive trades can contain a whole bid-ask spread (if one trade occurs at the bid and the other at the offer) and one minute of fundamental volatility, while a one-day return can also contain a whole bid-ask spread (if one close occurs at the bid and the other at the offer) but a much larger 24 hours of fundamental volatility. Note, however, that bid-ask bounce is only one of several sources of noise in prices.

Another point of comparison is with the noise induced in daily returns by "price pressure," that is, temporary deviations from efficient prices due to risk-averse liquidity providers being unwilling to provide unlimited liquidity. Recently, Hendershott and Menkveld (2014), using data on New York Stock Exchange (NYSE) intermediaries, estimate that at daily frequencies the distortions in midquote prices caused by price pressure (i.e., separate from the effect of bid-ask-bounce) are economically large (0.49% on average) and have a half-life of 0.92 days. The ratio of price pressure (in the midquote) to the variance of the efficient midquote price is 0.33 or 33% in their sample of 697 NYSE stocks during 1994-2005. This ratio of one source of noise to the estimated efficient price volatility is similar in magnitude to the estimated noise share of variance in our model.

Similarly, but at monthly frequencies and using a different approach, Hendershott et al. (2011) estimate that one-quarter of monthly return variance in NYSE stocks is due to transitory price changes that are themselves partially explained by cumulative order imbalances and market-makers' inventories (price pressure). Again, this is just one source of noise and in monthly returns, but it is also close in magnitude to our estimate.

Finally, our finding that a considerable proportion of the variance in daily returns is noise is consistent with studies such as Jegadeesh (1990) and Lehmann (1990) who document significant predictability (reversals) in stock returns at one month and one week horizons, respectively. Avramov, Chordia, and Goyal (2006) and Nagel (2012) show that the reversals reflect deviations from efficient prices. They find that non-informational demand generates price pressure that is reversed once liquidity suppliers react to potential profit opportunities and the uninformed demand for liquidity abates. While it is difficult to express the reversals documented by these studies as a percentage of variance to directly compare them to our estimates of noise, Jegadeesh (1990), Lehmann (1990), Avramov et al. (2006), and Nagel (2012) show that the price distortions involved in reversals are economically meaningful, consistent with the economically meaningful noise share estimated by our model. Similarly, Asparouhova et al. (2013) show that noise at daily frequencies causes an economically meaningful bias in returns, equal to 50% or more of the corrected estimate.

3.4. Variance components through time

Figure 3 shows how the stock return variance components change through time from 1960 to 2015. There are several noteworthy long-term trends. First, the amount of noise in prices has declined from around 40% of variance in the 1960s to around 20% of variance recently, although not monotonically. Noise rose through the 1990s, spiking in 1997, and has gradually declined since then. Table 2 Panel B confirms that stock returns after 1997 tend to have less noise and thus higher information content. The differences between the two sub-periods are statistically significant as well as economically meaningful. For example, the average noise share decreases from 35.47% before 1997 to only 25.69% after 1997.

The high levels of noise in prices in the 1990s are at least partly driven by collusive behavior of dealers during that period, which involved effectively widening the tick size by avoiding odd-eighth quotes and thereby increasing bid-ask bounce (Christie and Schultz, 1994). The post-1997 decline in noise to less than half of its peak levels is partly due to reductions in tick sizes starting in June 1997 as well as general improvements in liquidity and increases in turnover during the last two decades (e.g., Chordia, Roll, and Subrahmanyam, 2011). We provide more detailed analyses of these effects in Section 4.1.

Insert Figure 3 About Here

Second, while noise has declined through time, firm-specific information has become an increasingly important component of stock return variance. Together, the two firm-specific information components have increased from around 50% of variance in the early 1960s to above 70% of stock return variance in recent years. Table 2 Panel B confirms that this increase in firm-specific information is also statistically significant. The general trend is consistent with increasing informational efficiency through time. Interestingly, while public and private firm-specific information contribute approximately equally to stock return variance in the early 1960s, these components diverge through time with publically available firm-specific information emerging as the dominant component accounting for around 40% of stock return variance in recent years. The shift to public firm-specific information is consistent with the objectives of a variety of regulations such as the Sarbanes Oxley Act (2002) and Regulation Fair Disclosure (2000) to increase both the quality and quantity of public disclosure by companies.

As a short aside, our estimates of the proportion of firm-specific information that is impounded in prices through trading (private information) compared to public information are similar to Hasbrouck's estimates of the role of trading in impounding new information. Using intraday data, Hasbrouck (1991b) estimates that in 1989 34.3% of the information in prices gets impounded via individual trades. Despite differences in model, sample, and frequency, using daily data we estimate the fraction in 1989 to be around 39.1%.[11]

Third, while market-wide information tends to spike during crises, at other times it is generally not a substantial driver of individual stock returns. Throughout the sample period, market-wide information accounts for around 5-15% of stock return variance.

The broad trends illustrated in Figure 3 shed some light on recent issues concerning the information content of prices. For example, the concern that the growth in indexing and passive investing in recent years, and corresponding decline in active funds management, might harm the amount of firm-specific information in prices is not supported by the general trends in Figure 3. Similarly, suggestions that the

[11] A further difference is that our estimate of 39.1% corresponds to the fraction of *firm-specific* information in prices that is impounded through trading, while the 34.3% corresponds to the fraction of *all* information impounded in prices.

increase in market model R^2 and stock correlations since the late 1990s reflects a deterioration in informational efficiency is also not supported by the data, which show declining noise levels and increasing dominance of informational components. We explore whether or not composition changes in the market contribute to some of the time-series trends in the next subsection after analyzing the cross-sectional variation.

3.5. Variance components in the cross-section of stocks

Table 2 Panels C and D report means of each of the variance components in size quartiles and industry groups with these groupings formed each year. The returns of large stocks tend to reflect more market-wide information and more private firm-specific information. The differences are particularly large for market-wide information, which accounts for 21.51% of variance in big stocks, but only 4.53% in small stocks. Large stocks also tend to have less noisy prices. Noise declines monotonically with size and the differences across stocks are large. For example, in small stocks, noise accounts for 36.09% of stock return variance, which is about twice that of big stocks at 16.45% of variance. The relatively low level of noise in large stocks is likely driven by a high level of liquidity, making their prices less susceptible to temporary deviations and price pressures. Panel D shows that there is considerably less variation across industry groups in what drives stock price movements than across size groups.

To test the cross-sectional determinants of the variance components in a multivariate setting, we estimate the following panel regressions of stock-year observations:

$$Share_{i,t} = \alpha + \gamma_1 D_t^{POST} + \gamma_2 lnP_{i,t} + \gamma_3 lnMC_{i,t} + \gamma_3 D_i^{Consumer} + \gamma_4 D_i^{Healthcare} + \gamma_5 D_i^{HiTech} + \gamma_6 D_i^{Manufact} + \varepsilon_{i,t} \ , \quad (11)$$

where $Share_{i,t}$ is one of the variance component shares ($MktInfoShare_{i,t}$, $PrivateInfoShare_{i,t}$, $PublicInfoShare_{i,t}$, $NoiseShare_{i,t}$), D_t^{POST} is an indicator variable that takes the value of one after 1997 and zero before, $lnP_{i,t}$ is the log stock price, and $lnMC_{i,t}$ is the stock's log market capitalization. The indicator variables $D_i^{Consumer}$, $D_i^{Healthcare}$, D_i^{HiTech}, and $D_i^{Manufact}$ indicate the firm's industry group (the Other Industry group is the omitted category).

Insert Table 3 About Here

The regression results in Table 3 generally confirm the observations from the univariate analysis. Stock returns in the 1997-2015 part of the sample tend to contain significantly less noise and more public firm-specific information, even after controlling for other factors. Therefore, the time-series changes in noise and in firm-specific information are not driven simply by firms becoming larger through time. The returns of large stocks and high priced stocks are significantly more affected by market-wide information. Large stocks also tend to have less noisy prices and reflect relatively more firm-specific private information, controlling for other factors. Among the five industry groups, stocks in the Healthcare and Hi Tech sectors tend to have the highest levels of private firm-specific information and lowest levels of noise.

Given that the composition of stock return variance differs in the cross-section, in particular by size and to a lesser extent by industry, we examine to what extent the time-series patterns in variance components are due to market composition changes. The mix of industries in the market has changed through time and listed stocks have tended to become larger through time. Therefore, in Figures 4 and 5 we repeat the exercise of plotting the time-series of variance components, but this time by size group and by industry group. We form the size groups with respect to thresholds ($100 million and $1 billion in 2010 dollars) that are inflation adjusted through time, rather than size quartiles so as to keep the size groups relatively comparable through time even as the market composition changes. Time-series trends in the variance components within size or industry groups are less susceptible to compositional changes than the pooled time-series.

Figure 4 shows that all size groups have a similar trend with respect to market-wide information, including the peaks during crises. Large stocks consistently reflect more market-wide information through time than smaller stocks. All size groups show remarkably similar trends in private firm-specific information, except for a period of temporary divergence in the 1990s. The increase in public firm-specific information through time is driven mainly by smaller stocks consistent with improvements in their disclosure. Finally, noise is consistently higher for smaller stocks and smaller stocks are largely responsible for the decline in noise through time, in particular since the mid-1990s.

Insert Figure 4 About Here

Figure 5 shows that the variance components in different industry groups display remarkably similar time-series trends. Not only are the long-run trends in the types of information and noise similar across the industry groups, but so too are many of the year-to-year fluctuations. This result indicates that the time-series trends are not driven by changing industry composition in the market. Furthermore, it indicates that much of the variation in the information and noise shares is systematic and not just an artifact of estimation error or random fluctuations. Recall that the variance decomposition is performed separately (independently) for each stock in each year. The commonality in the variance component trends across groups of stocks (in this case industry groups) points to systematic drivers of the type of information and degree of noise in prices.

Insert Figure 5 About Here

4. Further validation tests

The preceding section provides some informal validation of the empirical variance decomposition model by showing that the variance components have reasonable time-series and cross-sectional properties, that they exhibit systematic variation, and that some of the estimated levels are consistent with estimates in other studies that use different approaches (usually capturing only one of the components). This section presents further validation tests that examine how the variance components respond to exogenous shocks to tick sizes (Section 4.1) and exogenous shocks to analyst coverage (Section 4.2). We also relate the variance components to other measures of noise and information (Section 4.3).

4.1. Exogenous shocks to tick sizes

Chordia, Roll, and Subrahmanyam (2008) show that decreases in tick sizes increase informational efficiency, implying a decrease in noise when tick sizes are smaller. During the sample period, the tick size was reduced from eighths of a dollar to sixteenths of a dollar on June 24, 1997, and then from sixteenths to pennies on January 29, 2001. This setting provides a natural experiment in which to test how the variance components, in particular noise, respond to the exogenous decreases in tick size and accompanying increase in informational efficiency.

If the estimated noise share of variance is indeed able to capture noise in prices we expect to see three patterns related to tick sizes and changes in tick sizes. First, stocks with larger relative tick sizes (tick size divided by price) should have noisier prices and thus a larger noise share. Given the tick size in dollars is the same for all stocks in the cross-section (at the time of the tick size changes) but price levels vary in the cross-section, we expect lower priced stocks to have higher levels of noise because they have larger relative tick sizes. Second, we expect that when the tick sizes are reduced, noise declines. Third, we expect that the effects of tick size reductions are heterogeneous in the cross-section with lower priced stocks having a larger decline in noise because for such stocks the change in the relative tick size is larger. For example, the tick size reduction from eighths of a dollar to sixteenths is 1.25% of the price of a $5 stock, but only 0.125% of the price of a $50 stock.[12]

To exploit this natural experiment, we take a subsample of one year on both sides of the tick size reduction from eighths of a dollar to sixteenths of a dollar (i.e., we take the years 1996 and 1998), and

[12] Similarly, the tick-to-price, a measure of the pricing grid coarseness, decreases by a larger amount for lower priced stocks.

estimate difference-in-differences models that exploit the cross-sectional heterogeneity in the treatment.[13] The highest priced quartile of stocks had the smallest change in relative tick size and therefore serves as a control group against which to measure the impact of the tick size reduction in other price quartiles:

$$\begin{aligned} NoiseShare_{it} = \alpha + \beta D^{POST}_{TickReduction,t} + \gamma_1 D^{POST}_{TickReduction,t} Q1_i + \gamma_2 D^{POST}_{TickReduction,t} Q2_i \\ + \gamma_3 D^{POST}_{TickReduction,t} Q3_i + \rho_1 Q1_i + \rho_2 Q2_i + \rho_3 Q3_i + \varepsilon_{it}, \end{aligned} \tag{12}$$

where $D^{POST}_{TickReduction,t}$ takes the value of one after the tick size reduction (1998) and zero otherwise. $Q1_i$, $Q2_i$, and $Q3_i$ are indicator variables that indicate the price quartile to which the firm belongs. The highest price quartile, $Q4_i$, is the omitted category. We also re-estimate the model in Equation (12) using the log price ($lnP_{i,t}$) instead of the price quartile indicators as a robustness test:

$$NoiseShare_{it} = \alpha + \beta D^{POST}_{TickReduction,t} + \gamma_1 D^{POST}_{TickReduction,t} lnP_{i,t} + \gamma_2 lnP_{i,t} + \varepsilon_{it}. \tag{13}$$

Insert Table 4 About Here

In Table 4, Models 1 and 2 show a monotonic relation between a stock's price level and its level of noise. As expected, lower priced stocks have higher levels of noise consistent with the fact that they have larger relative tick sizes (tick size divided by price). The difference in the cross-section of stocks is economically meaningful: the highest priced quartile has a noise share of around 19%, while the next three price quartiles have noise shares that are 6%, 8%, and 14% higher, with these differences being statistically significant. Models 3 and 4 show that noise tends to decline when the tick size is reduced consistent with our expectations. Furthermore, the decline in noise is larger for lower priced stocks consistent with the fact

[13] Chordia, Roll, and Subrahmanyam (2008) show that informational efficiency improved significantly around the change from eighths of a dollar to sixteenths and to a lesser extent from sixteenths to pennies.

that for such stocks the change in the relative tick size is larger. Therefore, the analysis of how noise relates to tick sizes or price discreteness and how it changes around exogenous changes in tick sizes support the notion that the noise share is a useful measure of the amount of noise in prices and returns.

We also examine a second natural experiment relating to the tick size. Christie and Schultz (1994) find evidence of collusive behavior by NASDAQ dealers during a period from 1991 until the collusive behavior was exposed a few years later. They show that NASDAQ dealers colluded to maintain artificially wide spreads by avoiding odd-eighth quotes. This behavior increased the effective tick size and due to bid-ask bounce is expected to increase the noise in prices. Importantly, a natural control group is non-NASDAQ stocks.

To examine this natural experiment, we take a subsample of four years before and during the collusive behavior and estimate a difference-in-differences model that uses non-NASDAQ stocks as a control group:

$$NoiseShare_{it} = \alpha + \beta D_t^{COLLUSION} + \gamma D_t^{COLLUSION} NASDAQ_i + \delta NASDAQ_i + \varepsilon_{it}, \qquad (14)$$

where $D_t^{COLLUSION}$ is an indicator variable that takes the value of one in the collusion period (1991-1994) and zero otherwise and the indicator variable $NASDAQ_i$ is an indicator variable that takes the value one for NASDAQ-listed stocks and zero otherwise.

Model 5 in Table 4 shows that during the period of collusion by NASDAQ dealers the returns of NASDAQ-listed stocks are significantly noisier, consistent with discreteness in price grids contributing to pricing errors and noise in returns. The magnitude of the effect is economically meaningful. The increase in noise for NASDAQ-listed stocks is estimated to be 8.32% of variance, which is large considering that the pooled sample mean noise share is around 30.78% of variance. The results therefore support the notion that collusion by NASDAQ dealers effectively widened the tick size and that the noise share from our variance decomposition model captures this increase in noise for the affected stocks.

4.2. Exogenous shocks to analyst coverage

Exogenous shocks to analyst coverage provide another natural experiment that changes the information environment for individual stocks. Given that analysts produce information about individual companies and disseminate this information to a variety of market participants, a reduction in analyst coverage is likely to reduce the amount of public firm-specific information in prices. As information in prices declines, the relative level of noise is likely to increase. Analyst coverage is expected to have little effect on market-wide information. The effects of analyst coverage on private firm-specific information is ambiguous: analyst-generated information that is made available to only some market participants might be impounded in prices through the course of those participants trading on the information (an increase in private information), but it might also crowd-out private information acquisition (a decrease in private information).

To test the impact of analyst coverage on the information and noise in prices, we use brokerage mergers/closures as a source of exogenous variations in analyst coverage. Broker mergers and closures are plausibly exogenous shocks because the termination of coverage is not driven by the characteristics or behavior of the firm (see Hong and Kacperczyk 2010; Kelly and Ljungqvist 2012; Brogaard et al., 2018). We obtain a list of broker mergers and closures that combines the lists from Hong and Kacperczyk (2010), spanning 1984 to 2005, and Kelly and Ljungqvist (2012), spanning 2000 to 2008. Combining these lists, merging with CRSP and IBES (Institutional Brokers' Estimate System) data, and imposing the requirement that both the acquirer and target brokers must provide overlapping coverage for at least one firm before the broker merger (as per Kelly and Ljungqvist, 2012) we have 41 mergers/closures of brokers that occur during the period 1989-2009. Using the mergers/closures data, we calculate the number of exogenous analyst disappearances per stock-year. These mergers and closures result in exogenous coverage shocks to 4,546 firm-year observations.

Using the exogenous analyst coverage shocks, we estimate the following difference-in-differences model:

$$Share_{i,t} = \gamma_i + \delta_t + \beta_1 CoverageShock_{i,t} + \varepsilon_{i,t}, \tag{15}$$

where $Share_{i,t}$ is one of the variance component shares ($MktInfoShare_{i,t}$, $PrivateInfoShare_{i,t}$, $PublicInfoShare_{i,t}, NoiseShare_{i,t}$) for stock i in year t, γ_i and δ_t are stock and time fixed effects, respectively, and $CoverageShock_{i,t}$ is the number of analyst disappearances due to mergers and closures of brokerage houses during the past two years.[14] We estimate the model above using the period from 1987 to 2011 given that the brokerage mergers and closures occur between 1989 and 2009 and we need to observe a two-year trend before and after the analyst disappearances.

The results in Table 5 show that exogenous decreases in analyst coverage are associated with a decline in public firm-specific information and an increase in the noise share of variance. These results are consistent with the notion that analysts produce firm-specific information that is made publically available and becomes reflected in prices. It also suggests the public firm-specific information component of variance from our variance decomposition model is able to detect this change in the information environment. The coefficient estimates indicate that the exogenous disappearance of each analyst is associated with a decline in public firm-specific information equal to around 1.59% of variance (for comparison, the pooled sample mean of public firm-specific information is around 37.11% of variance).

Shocks to analyst coverage have no significant effect on the amount of market-wide information in prices as expected. Neither do they have a significant impact on the amount of private firm-specific information in prices.

Insert Table 5 About Here

[14] In contrast to standard difference-in-differences models, here the "treatment" can have different magnitudes depending on how many analysts cease their coverage of a given stock. If the number of analyst disappearances is different in year $t - 1$ and year $t - 2$, we take the maximum of these two values.

4.3. Relation between variance components and other measures of information and noise

In the next two validation tests, we examine the relation between the variance components and two other measures of information in prices: the Hou and Moskowitz (2005) delay metric and variance ratios as per French and Roll (1986) and Chordia, Roll, and Subrahmanyam (2011).

The Hou and Moskowitz (2005) delay metric is a measure of how efficiently market-wide information is reflected in individual stock prices. It is constructed in each stock-year by estimating a regression of daily stock returns ($r_{i,t}$) on daily market returns ($r_{m.t}$) and ten lags of daily market returns:

$$r_{i,t} = \alpha_i + \beta_i r_{m.t} + \sum_k^{10} \delta_{i,k} r_{m,t-k} + \varepsilon_{i,t}. \tag{16}$$

Regression (16) is estimated once with all of the lags of market returns (unconstrained) to capture the unconstrained regression R^2 ($R^2_{unconstrained}$) and once without the lags of market returns (constraining all $\delta_{i,k}$ to zero) to capture the constrained regression R^2 ($R^2_{constrained}$). The delay metric is constructed from the two R^2 as follows to measure the incremental explanatory power of the lagged market returns:

$$Delay_{i,t} = 1 - \frac{R^2_{constrained}}{R^2_{unconstrained}}. \tag{17}$$

If market-wide information is perfectly and instantly reflected in the stock's prices, the two R^2s are equal and $Delay_{i,t} = 0$, but if stock prices are sluggish in reflecting market-wide information, then $R^2_{unconstrained} > R^2_{constrained}$ and $Delay_{i,t} > 0$. We therefore expect that higher values of $Delay_{i,t}$ should be associated with less market-wide information in prices (lower $MktInfoShare$) and less efficient, noisier prices (higher $NoiseShare$).

We estimate the relation between each of the variance components ($Share_{i,t}$) and the delay metric using the following panel regressions, including stock and year fixed effects (γ_i and δ_t):

$$Share_{i,t} = \gamma_i + \delta_t + \beta_1 Delay_{i,t} + \varepsilon_{i,t}. \quad (18)$$

Table 6 Panel A reports the result. There is a very strong inverse relation between the delay metric and the share of variance that is attributable to market-wide information by our variance decomposition. An increase in $Delay_{i,t}$ from zero (full efficiency) to 0.5 (half way to the maximum value of $Delay_{i,t}$) is associated with a reduction of $MktInfoShare$ by 9.17% of variance (a large effect, considering the pooled sample mean of $MktInfoShare$ is 8.24%. The results therefore indicate that $MktInfoShare$ reflects the efficiency with which market-wide information is reflected in prices.

.

Insert Table 6 About Here

Next we examine the variance ratios measure (French and Roll, 1986; Chordia et al., 2011), defined as the variance of returns during trading hours (variance of open-to-close returns, ${}^{1}/{}_{n}\sum_{1}^{n} r_{intraday,i,d}^{2}$) divided by the variance of overnight returns (variance of close-to-open returns, ${}^{1}/{}_{n}\sum_{1}^{n} r_{overnight,i,d}^{2}$). We measure the variance ratio for each stock i in each year t using daily, d, observations, excluding weekends:

$$VarianceRatio_{i,t} = \frac{{}^{1}/{}_{n}\sum_{1}^{n} r_{intraday,i,d}^{2}}{{}^{1}/{}_{n}\sum_{1}^{n} r_{overnight,i,d}^{2}}. \quad (19)$$

French and Roll (1986) and Chordia et al. (2011) show that the variance ratio is related to the amount of information reflected in prices, in particular through trading, and they use this ratio as a measure of informational efficiency. Under this interpretation, we expect the variance ratio to be positively related to most or all of the information components of variance and negatively related to the amount of noise in prices.

We estimate the relation between each of the variance components and the variance ratio metric using similar panel regressions as in Equation (18) including stock and year fixed effects. The results in Table 6 Panel B show that this is the case. The variance ratio is positively related to all of the information components of variance and negatively related to the amount of noise in prices. This result supports the interpretation of the variance ratio as an efficiency measure as well as the ability of the variance decomposition to separate out information from noise.

5. Extensions incorporating cash flow and discount rate news

This section extends the variance decomposition by separating each of the information components of variance into cash flow and discount rate parts. One reason for doing so is that by accounting for noise, decompositions of cash flow / discount rate news can be performed at higher frequencies (traditionally, monthly returns are used to minimize concerns about noise), which allows examination of the time-series trends in those information components.

First we review the standard approach for separating cash flow and discount rate news, developed by Campbell and Shiller (1988a, 1988b) and Campbell (1991) and subsequently used in many papers (Section 5.1). We then extend the standard approach by accounting for noise, noting how noise impacts the estimated cash flow and discount rate news (Section 5.2). Finally, we use cash flow / discount rate decompositions to produce an extended version of our variance decomposition (Section 5.3).

5.1. The standard approach to separating cash flow and discount rate news

Campbell and Shiller (1988a, 1988b) and Campbell (1991) show, without having to make behavioral or preference assumptions, that an unexpected stock return, $\varepsilon_{r_{t+1}}$, is equal to two parts:

$$\begin{aligned}\varepsilon_{r_{t+1}} &= (\mathrm{E}_{t+1} - \mathrm{E}_t)\sum_{j=0}^{\infty}\rho^j \Delta d_{t+1+j} - (\mathrm{E}_{t+1} - \mathrm{E}_t)\sum_{j=1}^{\infty}\rho^j r_{t+1+j} \\ &= \varepsilon_{CASHFLOWS_{t+1}} + \varepsilon_{DISCOUNT_{t+1}}, \end{aligned} \tag{20}$$

where $\varepsilon_{CASHFLOWS_{t+1}} = (\mathrm{E}_{t+1} - \mathrm{E}_t)\sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j}$ is cash flow news and $\varepsilon_{DISCOUNT_{t+1}} = -(\mathrm{E}_{t+1} - \mathrm{E}_t)\sum_{j=1}^{\infty} \rho^j r_{t+1+j}$ is discount rate news, d_t is the log dividend at time t, r_t is the log holding period return at time t, and $\rho \approx 0.96$ is a constant.

The terms in Equation (20) can be estimated from a VAR in which one of the variables is the log stock return.[15] The typical approach is to use the VAR to estimate discount rate news because that does not require information on dividends, and then obtain the cash flow news as the difference between the unexpected stock return and the discount rate news, $\varepsilon_{CASHFLOWS_{t+1}} = \varepsilon_{r_{t+1}} - \varepsilon_{DISCOUNT_{t+1}}$. The importance of cash flow news and discount rate news can be quantified by the variance or standard deviation of the two time-series: $\varepsilon_{CASHFLOWS_t}$ and $\varepsilon_{DISCOUNT_t}$.

5.2. Accounting for noise when separating cash flow and discount rate news

A limitation of the standard approach (summarized above) for separating cash flow and discount rate news is that it does not account for the noise in stock returns. Without accounting for noise, the cash flow / discount rate decomposition can only be reliably performed using low-frequency data such as monthly returns so that the ratio of noise to information remains within acceptable error tolerances. Therefore, the standard approach is limited in its ability to examine time-series variation in the cash flow / discount rate components. For example, with monthly returns and a minimum of 20 time-series observations in the VAR, one can obtain a single value of cash flow and discount rate variance every ten years. Accounting for noise, however, allows us to apply the decomposition to daily data and thereby estimate cash flow and discount rate news variances every year. This higher resolution reveals time-series trends in cash flow and discount rate news and also enables us to further partition the information components in our baseline model.

[15] For example, once the VAR is estimated, one can obtain the time t expectations of returns at $t+2$, $t+3$ and so on (multi-step forecasts from the VAR) from which one can compute $\sum_{j=1}^{\infty} \rho^j \mathrm{E}_t[r_{t+1+j}]$. Repeating this process at time $t+1$ one obtains $\sum_{j=1}^{\infty} \rho^j \mathrm{E}_{t+1}[r_{t+1+j}]$. The difference gives the discount rate news at time $t+1$, i.e., $\varepsilon_{DISCOUNT_{t+1}} = -(\mathrm{E}_{t+1} - \mathrm{E}_t)\sum_{j=1}^{\infty} \rho^j r_{t+1+j} = \sum_{j=1}^{\infty} \rho^j \mathrm{E}_t\left[r_{t+1+j}\right] - \sum_{j=1}^{\infty} \rho^j \mathrm{E}_{t+1}[r_{t+1+j}]$.

To understand how noise manifests in a standard cash flow / discount rate decomposition and therefore how to approach the task of isolating noise in the decomposition, consider Figure 6 Panel A. A stock return is composed of a discount rate that captures the required or expected rate of return, noise, and information. Noise has an expected and an unexpected component. The expected component arises from reversals of pricing errors. For example, a positive pricing error is expected to reverse resulting in an expected negative return component.[16] The unexpected component of noise reflects random changes to the pricing errors. Thus, the expected return is made up of the discount rate and the return from the expected change in the pricing error.

Insert Figure 6 About Here

The unexpected return is driven by information arrivals and shocks to pricing errors (unexpected noise). Therefore, in the standard cash flow / discount rate decomposition, noise contaminates the estimated discount rate news because the expected return reflects the discount rate *and* noise. Noise also contaminates the estimated cash flow news component because: (i) cash flow news is usually calculated as the difference between the unexpected stock return and the discount rate news, which is contaminated by noise; and (ii)

[16] There are several reasons why pricing errors can be inferred from past returns and their reversals are somewhat predictable. At the most basic level, bid-ask bounce (trade prices oscillating between the bid and the ask or offer quotes) creates negative serial correlation in returns and therefore a predictable "noise" component of returns (e.g., Roll, 1984). For example, if a stock's closing price is at the bid quote, its next close could be at the bid or the ask/offer and therefore, merely on the basis of bid-ask bounce, in expectation the next closing price will be higher, i.e., there is an expected positive noise return. Return predictability due to pricing errors goes well beyond the bid-ask bounce effect. Negative serial correlation is also found in midquote returns of individual stocks (suggesting quoted prices also suffer from temporary mean-reverting pricing errors) and at longer horizons such as weekly and monthly returns (e.g., Jegadeesh, 1990; Lehmann, 1990; Hendershott and Menkveld, 2014). The economically meaningful reversals in returns at daily through to monthly horizons (which has been validated in many studies) is linked to imperfect liquidity and the inability for the market to absorb order imbalances without temporarily deviating from efficient prices (e.g., Avramov et al., 2006; Hendershott et al., 2011; Nagel, 2012). The existence of predictable reversals in returns due to temporary price distortions from efficient prices is also supported by market microstructure theory. For example, classic inventory control models of liquidity supply show that when risk averse liquidity providers receive many buy orders, they "shade" their subsequent quoted prices upward (above the efficient price) to attract sellers and thereby revert their inventory towards zero, and vice versa when they receive many sell orders (e.g., Stoll, 1978, Ho and Stoll, 1981, and many subsequent models). Similarly, return reversals due to distortions from efficient prices arise in models of adverse selection (e.g., Kyle, 1985; Glosten and Milgrom, 1985) if liquidity providers are risk averse or less than perfectly competitive (e.g., Subrahmanyam, 1991; Nagel, 2012).

part of the unexpected return, which goes into the cash flow news calculation, is noise. To resolve these issues, our modified cash flow / discount rate decomposition first removes noise from both the expected and unexpected returns, resulting in a method that is suitable for higher frequency data.

First we modify our baseline model to allow for a time-varying discount rate. The efficient price from Equation (3) becomes:

$$m_t = m_{t-1} + \mu_t + w_t \ , \tag{21}$$

and the stock return from Equation (4) becomes:

$$r_t = p_t - p_{t-1} = \mu_t + w_t + \Delta s_t \ , \tag{22}$$

where the time-varying drift, μ_t, is the discount rate on the stock over the time t period, w_t is an innovation that reflects new information about the stock's fundamentals, and Δs_t is the change in pricing error. Noise has an expected component ($\mathrm{E}_{t-1}[\Delta s_t]$) and an unexpected component (ε_{s_t}), $\Delta s_t = \mathrm{E}_{t-1}[\Delta s_t] + \varepsilon_{s_t}$. The expected component comes from the fact that pricing errors are temporary and therefore tend to reverse, as discussed above. Consequently, the expected return ($\mathrm{E}_{t-1}[r_t]$) is made up of the discount rate and the expected change in the pricing error, $\mathrm{E}_{t-1}[r_t] = \mu_t + \mathrm{E}_{t-1}[\Delta s_t]$. Similarly, the unexpected return ($\varepsilon_{r_t} = r_t - \mathrm{E}_{t-1}[r_t]$) is made up of new information about the stock's fundamentals and unexpected changes in the pricing error (noise), $\varepsilon_{r_t} = w_t + \varepsilon_{s_t}$.

The information-driven innovation in the efficient price is the same as in our baseline model and is estimated from the VAR/VMA: $w_t = \theta_{r_m}\varepsilon_{r_m,,t} + \theta_x\varepsilon_{x,t} + \theta_r\varepsilon_{r,t}$. We estimate the expected return on the stock over the next period, $\mathrm{E}_{t-1}[r_t]$, as the one-period-ahead forecast of the return from the VAR, in the spirit of Campbell (1991). We isolate the expected noise part of the expected return by considering what part of the expected return is predicted by past unexpected changes in the pricing error, $\mathrm{E}_{t-1}[\Delta s_t] =$

$\frac{Cov(\mathrm{E}_{t-1}[r_t],\varepsilon_{s_{t-1}})}{Var(\varepsilon_{s_{t-1}})}\varepsilon_{s_{t-1}}$.[17] The other part of the expected return is the discount rate, $\mu_t = \mathrm{E}_{t-1}[r_t] - \mathrm{E}_{t-1}[\Delta s_t]$. We obtain the unexpected innovations in the pricing error from $\varepsilon_{s_t} = r_t - \mathrm{E}_{t-1}[r_t] - w_t = r_t - \mathrm{E}_{t-1}[\Delta s_t] - \mu_t - w_t$ and consequently the total change in the pricing error (sum of expected and unexpected parts) is $\Delta s_t = \mathrm{E}_{t-1}[\Delta s_t] + \varepsilon_{s_t} = r_t - \mu_t - w_t$.

A simple schematic of what is going on in the process above is shown in Figure 6. We break noise into expected and unexpected parts. Subtracting expected noise from the expected return gives the "clean" discount rate. The clean discount rate is similar to the discount rate in Campbell (1991) but purged of noise. Subtracting unexpected noise from the unexpected return gives the "clean" information. The clean information is similar to the cash flow and discount rate information in Campbell (1991) but purged of noise.

Next we apply a cash flow / discount rate decomposition similar to Campbell (1991), but using the clean discount rate and the clean information. Using the de-noised expected return ($\mathrm{E}_t[\mu_{t+1}]$) in place of the standard expected return ($\mathrm{E}_t[r_{t+1}]$), we estimate discount rate news using the Campbell (1991) approach:

$$\begin{aligned}\varepsilon_{DISCOUNT_{t+1}} &= -(\mathrm{E}_{t+1} - \mathrm{E}_t)\sum_{j=1}^{\infty}\rho^j \mu_{t+1+j} \\ &= \sum_{j=1}^{\infty}\rho^j \mathrm{E}_t[\mu_{t+1+j}] - \sum_{j=1}^{\infty}\rho^j \mathrm{E}_{t+1}[\mu_{t+1+j}]. \end{aligned} \tag{23}$$

Also following Campbell (1991), but using the de-noised unexpected return instead of the standard unexpected return, we estimate the cash flow news at time $t+1$ as the informational part of the return that is not associated with discount rate news:

[17] This approach is equivalent to estimating the predictive regression, $\mathrm{E}_{t-1}[r_t] = a + b\varepsilon_{s_{t-1}} + e_{t-1}$, where the estimate of the coefficient b is given by $\hat{b} = \frac{Cov(\mathrm{E}_{t-1}[r_t],\varepsilon_{s_{t-1}})}{Var(\varepsilon_{s_{t-1}})}$ and the part of $\mathrm{E}_{t-1}[r_t]$ that is explained by $\varepsilon_{s_{t-1}}$ is $\hat{b}\varepsilon_{s_{t-1}}$. This approach picks up the first-order negative serial correlation in returns that occurs at daily frequencies due to bid-ask bounce and price pressures.

$$\varepsilon_{CASHFLOWS_{t+1}} = w_{t+1} - \varepsilon_{DISCOUNT_{t+1}}. \tag{24}$$

From the time-series of the cash flow and discount rate news, we compute the variances $Var(\varepsilon_{CASHFLOWS_t})$ and $Var(\varepsilon_{DISCOUNT_t})$. We also compute the variance of noise, $Var(\Delta s_t)$.[18] We then plot the cash flow news, the discount rate news, and the noise as shares of variance.

Figure 7 plots the time-series of the cash flow news, discount rate news, and noise, expressed as shares of stock return variance.[19]

Insert Figure 7 About Here

Panel A reports results from the standard model that does not account for noise as represented in Equation (20), while Panel B is the model that accounts for noise and is described in Equations (23) and (24). In the model that does not account for noise, cash flow news is estimated to account for around 75% of stock return variance, while discount rate information makes up around 10%. The remaining variation is attributable to time-series variation in the discount rate itself (15%), which is different from discount rate

[18] The variance of noise differs slightly from our baseline model because we allow for a time-varying discount rate.

[19] In expressing the variance components as "shares" of variance, to make the results comparable to other models in the paper, we must also consider the covariance between cash flow and discount rate news. Given the total information in this model is the same as in the baseline model, to ensure the sum of the information component variances in this model equal the variance of information in the baseline model, we allocate a fraction α of $2Cov(\varepsilon_{DISCOUNT_t}, \varepsilon_{CASHFLOWS_t})$ to the cash flow news variance and a fraction $(1-\alpha)$ to the discount rate news variance, where $\alpha = \frac{Var(\varepsilon_{DISCOUNT_t})}{Var(\varepsilon_{DISCOUNT_t})+Var(\varepsilon_{CASHFLOWS_t})}$. Doing so does not change the ratio of cash flow news to discount rate news and, for consistency, we apply this covariance attribution to both the models that account for noise and those that do not.

news.[20] These results are consistent with Vuolteenaho (2002) who also performs a variance decomposition on individual stocks without accounting for noise and finds similar estimates.[21]

Other studies have performed similar decompositions on portfolios of stocks rather than individual stocks (e.g., Campbell, 1991; Campbell and Ammer, 1993). In portfolios, discount rate news plays a larger role, suggesting that cash flow news is more idiosyncratic than discount rate news. The dominance of cash flow information in our stock-level variance decomposition and the fact that cash flow information tends to be relatively idiosyncratic is also consistent with our baseline decomposition, which shows that idiosyncratic information is a far more important driver of individual stock returns than market-wide information.

Figure 7, Panel B adjusts the standard cash flow / discount rate decomposition for noise and reveals some interesting differences. A striking result is that almost all of the stock price variations associated with information is driven by cash flow news, with very little variation attributed to discount rate news. In fact, cash flow news is responsible for 68% of stock return variance in the full sample, whereas discount rate news accounts for less than 3%. It is natural to expect that accounting for noise would decrease both of these information components as some of the variation labelled as information in the standard models is noise. The interesting observation is that they do not decrease by a similar amount. The decrease in estimated discount rate news is far greater, resulting in a substantial increase in the estimated ratio of cash flow news to discount rate news when accounting for noise.

The results suggest that much of what is usually labelled as discount rate news is actually noise. Why? The primary reason is that noise creates considerable return predictability, so expected returns are

[20] The time-varying discount rate, $\mathrm{E}_t[r_{t+1}]$ in the model that does not account for noise and μ_t in the model that does account for noise, gives rise to variation in returns directly by determining the average rates of return in different periods, whereas the discount rate *news* captures price changes that occur when expectations of the discount rate change and the stock is re-priced accordingly. Given our focus on information and noise, we do not report the time-varying discount rate variance share in the plots.

[21] To better compare with Vuolteenaho (2002), we also calculate the ratio of cash flow news variance to discount rate news variance over the period from 1960 to 1996. Despite differences in data frequency and VAR model used, the ratio of cash flow news variance to discount rate news variance is about five times in our model, which is very similar to the ratio reported in Vuolteenaho (2002) for the same period of time.

not good measures of discount rates. Noise creates return predictability because pricing errors are stationary, mean reverting processes. Prices are drawn towards fundamental values in the long run, so a positive noise-driven return shock in one period leads to a negative expected return component over the next period and vice versa. The empirical consequence of pricing error reversals is the widely documented negative serial correlation in returns, which is observed at a wide range of frequencies from the classic monthly reversals anomaly (e.g., Jegadeesh, 1990) to weekly, daily, and intraday horizons (e.g., Roll, 1984). Without accounting for noise, variation in the discount rate is overestimated when the expected/forecast return is taken as an estimate of the discount rate, leading to a substantial overestimation of the discount rate news component.

Estimates of cash flow news are also affected by explicitly accounting for noise, but to a lesser extent due to two opposing effects. These effects are best illustrated by recognizing that cash flow news is the difference between estimated information and estimated discount rate news: $\varepsilon_{CASHFLOWS_t} = w_t - \varepsilon_{DISCOUNT_t}$. First, removing noise shrinks the estimated information shocks (w_t), which tends to decrease cash flow news. But, second, as explained above, the estimated discount rate news ($\varepsilon_{DISCOUNT_t}$) is considerably smaller after accounting for noise and this effect tends to increase the estimated cash flow news. The opposing effects explain why the estimated magnitude of cash flow news is less affected by accounting for noise than is the estimated magnitude of discount rate news.

An advantage of isolating noise is the ability to apply the decomposition over relatively short windows using high-frequency data. Unlike previous studies, this allows us to examine the time-series variation in the cash flow and discount rate news. Figure 7 shows that since the late 1990s, there has been a notable increase in the proportion of stock returns that are attributable to cash flow news, mirroring the decrease in noise during the same period. This trend matches our earlier decomposition that shows firm-specific information has become an increasingly important component of stock returns during the past two decades, consistent with the widely held view that financial markets are now more informationally efficient than in previous decades.

5.3. Extended variance decomposition

Armed with a method to separate cash flow and discount rate news at the daily frequency purged of noise, we further extend our baseline variance decomposition by splitting each information component into a cash flow part and a discount rate part. This extended decomposition of information is illustrated in Figure 5 Panel B. Note that the noise and time-varying discount rate components are not shown.

The six information components in the extended decomposition are obtained from the following regressions of cash flow and discount rate news on each of the information components from our variance decomposition:

$$\varepsilon_{DISCOUNT_t} = \beta_1 r_{A,t} + \beta_2 r_{B,t} + \beta_3 r_{C,t}$$

$$\varepsilon_{CASHFLOWS_t} = \gamma_1 r_{A,t} + \gamma_2 r_{B,t} + \gamma_3 r_{C,t}, \tag{25}$$

where the information components are market-wide information ($r_{A,t} = \theta_{r_m}\varepsilon_{r_m,t}$), firm-specific private information ($r_{B,t} = \theta_x \varepsilon_{x,t}$), and firm-specific public information ($r_{C,t} = \theta_r \varepsilon_{r,t}$).[22] From the fitted values we obtain six sources of variance: market-wide discount rate and cash flow news, $\widehat{\beta_1} r_{A,t}$ and $\widehat{\gamma_1} r_{A,t}$, firm-specific discount rate and cash flow news incorporated through trading on private information, $\widehat{\beta_2} r_{B,t}$ and $\widehat{\gamma_2} r_{B,t}$, and firm-specific discount rate and cash flow news incorporated through public information, $\widehat{\beta_3} r_{C,t}$ and $\widehat{\gamma_3} r_{C,t}$, respectively. In expressing the variance components as variance shares, we add back the covariance between cash flow and discount rate news as before, preserving the total variance attributable to information.

Insert Table 7 About Here

[22] In the regression, $\beta_i + \gamma_i = 1$, thereby preserving the total amount of each information type.

Table 7 reports means of the seven variance components from our extended decomposition, namely six information components and a noise component, expressed as percentages of variance.[23] The pooled sample results are presented in Panel A. Panel B shows the results separately for the two subperiods, before and after 1997. Results for size, price, and industry subgroups are presented in Panel C, D, and E, respectively. Consistent with our earlier observation corroborating Chen et al. (2013) that cash flow news is a much larger driver of individual stock returns than discount rate news, we also find that the cash flow parts of the market-wide and firm-specific information components are much larger than the corresponding discount rate parts. Overall, firm-specific cash flow information comprises the largest contribution to individual stock return variance, accounting for 27% of variance (the sum of the CF columns for $PrivateInfoShare$ and $PublicInfoShare$ in Table 7 Panel A).

What is perhaps more interesting is that the ratio of cash flow to discount rate news differs across the three information components. The differences are consistent with the notion that cash flow news tends to be more idiosyncratic than discount rate news. For example, the ratio of cash flow news to discount rate news in the firm-specific information component is around 29 times, whereas in market-wide information it is around 18 times. We observe this relation in all price and size quartiles as well as industry groups. This finding helps reconcile differing results in the literature: when variance decompositions are performed on portfolios of stocks (e.g., Campbell, 1991; Campbell and Ammer, 1993), in which most of the firm-specific variation is cancelled out through diversification, leaving predominantly market-wide information, discount rate news tends to be more important than when variance decompositions are performed on individual stocks (e.g., Vuolteenaho, 2002; Chen et al., 2013). We extend these findings by showing that once we isolate the noise component of returns, the importance of cash flow news relative to discount rate news becomes even more apparent.

[23] For conciseness, we do not report the share of variance attributable to time-variation in the discount rate (μ_t), which is why the seven reported components sum to slightly less than 100%.

6. Conclusion

This study decomposes stock return variance in order to better understand the roles of different types of information and noise in driving stock price movements. We find that a substantial proportion of return variance, 31%, is noise. Firm-specific information accounts for the majority (61%) of stock return variance, with market-wide information accounting for the remaining 8% of variance in the full sample. We further partition firm-specific information and find that in the full sample, public firm-specific information plays a larger role than private firm-specific information that is impounded into prices through trading.

We also find that after accounting for the noise in returns, cash flow information is more important than the previous literature suggests. Cash flow information plays a considerably larger role than discount rate information in driving individual stock returns. Discount rate information plays a relatively larger role in market-wide information than it does in information about individual firms.

There is substantial time-series variation in the components of variance, with some key trends standing out. First, noise increases from the 1970s to the mid-1990s, in particular around a period of collusion by dealers to effectively widen bid-ask spreads, and has substantially declined since then. The decline in noise is attributable in part to narrower tick sizes, which reduces bid-ask bounce, and a general improvement in liquidity and increase in turnover. We show that the recent decrease in noise due to improved liquidity is largely responsible for the increasing R^2 of a market model over the past two decades. An important implication is that a lower R^2 is not necessarily associated with more informationally efficient prices, in contrast to the interpretation of R^2 in prior studies.

Second, the role of firm-specific information has increased through time, driven largely by increases in the amount of public firm-specific information that is reflected in prices. This trend is consistent with increasing informational efficiency through time. The increasing importance of public firm-specific information in stock prices is also consistent with a variety of regulatory reforms such as the Regulation Fair Disclosure (2000) and the Sarbanes Oxley Act (2002) aimed at improving both the quality and quantity of corporate disclosure.

Third, market-wide information has over time become a less important driver of stock returns. While market-wide information tends to spike during crises it has generally declined from around 15% of variance to around 5-10% in recent years.

Overall, the broad trends in the components of stock return variance shed light on recent issues concerning the information content of prices. For example, the concern that the growth in indexing and passive investing in recent years and corresponding decline in active fund management might harm the amount of firm-specific information in prices is not intuitively supported by the observed time trends. Similarly, suggestions that the increase in the market model R^2 and stock correlations in recent years reflects a deterioration in informational efficiency is also not supported by the data, which show declining noise levels and increasing informational components.

While our results provide some new insights about these general issues, we leave a more detailed examination of each of these issues to future research. This paper's contribution is largely methodological. The framework for variance decomposition developed in this paper can be applied to analyzing each of these issues and others, due to (i) its ability to isolate noise from information, which is crucial for correctly characterizing the information in prices, and (ii) the ability to obtain higher frequency estimates of variance components, which is important in analyzing effects that vary through time and recent phenomena that require high resolution estimates of the information/noise components of variance.

Appendix A: Motivating theoretical model

A.1. Model

The model starts with the firm's cash flow generating process:

$$C_t = K_0 X_t, \tag{A.1}$$

where C_t is the future cash flow at time t, K_0 is the initial investment, and X_t captures the random shocks to the cash flow process. Departing from the Jin and Myers (2006) model, in which the focus is solely on information innovations, we suppose that X_t, as it is perceived or estimated by investors ($\widetilde{X_t}$), is driven by four components, namely three information-driven components and noise:

$$\widetilde{X_t} = \theta_{1,t} + \theta_{2,t} + \theta_{3,t} + \theta_{4,t}. \tag{A.2}$$

The term $\theta_{1,t}$ reflects market-wide cash flow information, while $\theta_{2,t}$ and $\theta_{3,t}$ reflect firm-specific cash flow information, and $\theta_{4,t}$ reflects noise. The distinction between $\theta_{2,t}$ and $\theta_{3,t}$ is that $\theta_{2,t}$ is firm-specific information revealed through trading on private information, while $\theta_{3,t}$ is firm-specific information revealed through public information such as company announcements and news. Firm-specific information ($\theta_{2,t}$ and $\theta_{3,t}$) is orthogonal to market-wide information ($\theta_{1,t}$) and noise is assumed to be uncorrelated with information.

The noise ($\theta_{4,t}$) is a reduced form way to account for many sources of noise in prices. It accounts for the fact that prices set by investors following their beliefs about $\widetilde{X_t}$ can deviate from intrinsic values that are based solely on true information. The deviations can arise from rational causes such as imperfect signals, or various friction, or irrational reasons and biases. The addition of $\theta_{4,t}$ to the model is in the spirit of Black's (1986) notion that "noise trading is trading on noise as if it were information", i.e., investors believe $\theta_{4,t}$ to be information about cash flows, causing prices to depart from their efficient intrinsic values. The ultimate effect of injecting noise into Equation (A.2) is to make prices noisy, which can be interpreted as representing the effects of frictions such as a discrete pricing grid, non-synchronous trading and so on.

Similar to Jin and Myers (2006), we assume that all of the information and noise components follow stationary AR(1) processes driven by a set of random shocks ($\varepsilon_{1,t}$, $\varepsilon_{2,t}$, $\varepsilon_{3,t}$, and $\varepsilon_{4,t}$).[24] That is, $\theta_{i,t+1} = \theta_{i,0} + \varphi\theta_{i,t} + \varepsilon_{i,t+1}$, where $0 < \varphi < 1$. We define r as the discount rate and K_t as the investors' valuation

[24] This and other assumptions made in this appendix are for the convenience of solving the motivating theoretical model (they mainly follow from Jin and Myers (2006)). We do not rely on these assumptions in the empirical model.

of the firm, which is the present value of future cash flows (assuming each period's cash flow is paid out), conditional on the information that the investors have at date t about $\widetilde{X_t}$:

$$K_t = PV\{E(C_{t+1}|I_t), E(C_{t+2}|I_t), \dots ; r\}. \tag{A.3}$$

For investors $I_t = \{\theta_{1,t} + \theta_{2,t} + \theta_{3,t} + \theta_{4,t}\}$ and therefore the value of K_t is:

$$K_t = \frac{K_0 X_0}{r(1-\varphi)} - \frac{K_0 X_0 \varphi}{(1+r-\varphi)(1-\varphi)} + \frac{\varphi}{1+r-\varphi} K_0(\theta_{1,t} + \theta_{2,t} + \theta_{3,t} + \theta_{4,t}). \tag{A.4}$$

Let $\tilde{r}_{t+1}$ be the total realized return on the firm's shares in the period $t+1$. The return $\tilde{r}_{t+1}$ is calculated as the change in the investors' valuation of the firm from one period to the next plus that period's cash flow, which is paid out, and is therefore a function of the shocks to the investors' information about the cash flow process:

$$\tilde{r}_{t+1} = r + b_t\left(\varepsilon_{1,t+1} + \varepsilon_{2,t+1} + \varepsilon_{3,t+1} + \varepsilon_{4,t+1}\right), \tag{A.5}$$

where

$$b_t = \frac{(1+r)}{\frac{X_0(1+r)}{r} + \varphi(\theta_{1,t} + \theta_{2,t} + \theta_{3,t} + \theta_{4,t})}. \tag{A.6}$$

The derivation of Equations (A.5) and (A.6) is provided below in Appendix A.2 (Proof 1). Equation (A.6) implies that the random component of realized stock returns, $b_t\left(\varepsilon_{1,t+1} + \varepsilon_{2,t+1} + \varepsilon_{3,t+1} + \varepsilon_{4,t+1}\right)$, is driven by shocks to the various types of information (market-wide information, firm-specific information revealed through trading on private information, and public firm-specific information) as well as innovations in noise.

The model suggests stock return variance can be decomposed into four distinct sources of variation. Performing a variance decomposition on the realized returns in this theoretical framework serves as a guide for the empirical decomposition. Because of the independence between the components of realized returns, the variance of realized returns is equal to the sum of the contributions from each of the components. This allows us to define a set of variance shares as:

$$\eta_j = \frac{Var(\varepsilon_{j,t})}{Var(\varepsilon_{1,t} + \varepsilon_{2,t} + \varepsilon_{3,t} + \varepsilon_{4,t})}, \tag{A.7}$$

with subscript $j = \{1, 2, 3, 4\}$ denoting each return component. The first of these variance shares, η_1, is the contribution of market-wide information to an individual stock's variance. The second (η_2) is the contribution of firm-specific information that is revealed through trading (private firm-specific information), while the third (η_3) is the contribution of public firm-specific information. The last component (η_4) is the effect of noise on stock return variance.

A.2. Proofs

Proof 1.

The total realized return on the firm's shares for outsiders is calculated as the change in the investors' valuation of the firm from one period to the next including the cash flow that is paid out:

$$\tilde{r}_{t+1} = \frac{K_{t+1} + C_{t+1}}{K_t} - 1. \tag{A.8}$$

Substituting Equation (A.4) for K_t into (A.8) and rearranging gives:

$$\tilde{r}_{t+1} = \frac{\left(\frac{1+r}{1+r-\varphi}\right) C_{t+1} - \frac{\varphi}{1+r-\varphi} C_t}{\frac{(1+r)}{r} \frac{K_0 X_0}{(1+r-\varphi)} + \frac{\varphi}{1+r-\varphi} C_t}. \tag{A.9}$$

For investors $I_t = \{\theta_{1,t} + \theta_{2,t} + \theta_{3,t} + \theta_{4,t}\}$ and therefore,

$$\tilde{r}_{t+1} = \frac{\left(\frac{1+r}{1+r-\varphi}\right) K_0 \sum_{j=1}^{4} \theta_{j,t+1} - \frac{\varphi}{1+r-\varphi} K_0 \sum_{j=1}^{4} \theta_{j,t}}{\frac{(1+r)}{r} \frac{K_0 X_0}{(1+r-\varphi)} + \frac{\varphi}{1+r-\varphi} K_0 \sum_{j=1}^{4} \theta_{j,t}}. \tag{A.10}$$

Multiplying the denominator and the numerator by $(\frac{1+r-\varphi}{K_0})$ and rearranging, we get:

$$\tilde{r}_{t+1} = r + \frac{(1+r)\left[\sum_{j=1}^{4}\theta_{j,t+1} - \varphi\sum_{j=1}^{4}\theta_{j,t} - X_0\right]}{\frac{(1+r)X_0}{r} + \varphi\sum_{j=1}^{4}\theta_{j,t}}. \tag{A.11}$$

Given the assumptions that all of the information and noise components follow stationary AR(1) process (we do not make this assumption in the empirical model), Equation (A.11) becomes:

$$\tilde{r}_{t+1} = r + b_t\left(\varepsilon_{1,t+1} + \varepsilon_{2,t+1} + \varepsilon_{3,t+1} + \varepsilon_{4,t+1}\right), \tag{A.12}$$

where

$$b_t = \frac{(1+r)}{\frac{(1+r)X_0}{r} + \varphi(\theta_{1,t} + \theta_{2,t} + \theta_{3,t} + \theta_{4,t})}. \tag{A.13}$$

Proof 2.

The realized return on stock i in Equation (A.12) can be rewritten as:

$$\tilde{r}_{i,t+1} = r + b_t\varepsilon_{1,t+1} + b_t\varepsilon_{f,t+1}, \tag{A.14}$$

where

$$b_t = \frac{(1+r)}{\frac{(1+r)X_0}{r} + \varphi(\theta_{1,t} + \theta_{2,t} + \theta_{3,t} + \theta_{4,t})},$$

and $\varepsilon_{f,t+1}$ is the shock related to firm-specific information and noise,

$$\varepsilon_{f,t+1} = \varepsilon_{2,t+1} + \varepsilon_{3,t+1} + \varepsilon_{4,t+1}. \tag{A.15}$$

The market return is the same as the return of a stock with no idiosyncratic risk:

$$\tilde{r}_{m,t+1} = r + d_t\varepsilon_{1,t+1}, \tag{A.16}$$

where

$$d_t = \frac{(1+r)}{\frac{(1+r)X_0}{r} + \varphi\theta_{1,t}}. \tag{A.17}$$

From Equation (A.16), we have:

$$\varepsilon_{1,t+1} = \frac{r_{m,t+1} - r}{d_t}. \tag{A.18}$$

Substitute the expression in Equation (A.18) for $\varepsilon_{1,t+1}$ into (A.14):

$$\tilde{r}_{i,t+1} = r + \frac{b_t}{d_t} r_{m,t+1} - \frac{b_t}{d_t} r + b_t \varepsilon_{f,t+1}. \tag{A.19}$$

Conditional on $\theta_{1,t}$, $\theta_{2,t}$, $\theta_{3,t}$ and $\theta_{4,t}$, the stock return variance is therefore a function of shocks to the investors' information about the cash flow process:

$$Var(\tilde{r}_{i,t+1}) = \left(\frac{b_t}{d_t}\right)^2 Var\left(r_{m,t+1}\right) + Var(b_t \varepsilon_{f,t+1}) \tag{A.20}$$

Substituting (A.16) into Equation (A.20), and rearranging, we get:

$$Var(\tilde{r}_{i,t+1}) = b_t^2\left[Var\left(\varepsilon_{1,t+1}\right) + Var(\varepsilon_{2,t+1}) + Var(\varepsilon_{3,t+1}) + Var(\varepsilon_{4,t+1})\right] \tag{A.21}$$

From Equation (A.19), the proportion of variance explained by the market, R^2, can be written as:

$$\begin{aligned} R^2 &= \frac{\left(\frac{b_t}{d_t}\right)^2 Var(r_{m,t+1})}{Var(\tilde{r}_{f,t+1})} \\ &= \frac{Var(\varepsilon_{1,t+1})}{Var\left(\varepsilon_{1,t+1}\right) + Var(\varepsilon_{2,t+1}) + Var(\varepsilon_{3,t+1}) + Var(\varepsilon_{4,t+1})} \\ &= \frac{1}{1 + \frac{1}{\eta_1}(\eta_2 + \eta_3 + \eta_4)}. \end{aligned} \tag{A.22}$$

Appendix B: Variable definitions

The table below provides descriptions and notation for the variables that are components of stock return variance. Each variable is estimated separately for each stock in each year using daily observations. When aggregating across stocks, we take variance-weighted averages (as per Morck et al., 2000, 2013).

Variable	Notation	Description
Stock return co-movement	R^2	R^2 is estimated by regressing individual daily stock returns on daily market return.
Noise share	$NoiseShare$	The share of stock return variance that is attributable to noise.
Market-wide information share	$MktInfoShare$	The share of stock return variance that is attributable to market-wide information.
Private firm-specific information share	$PrivateInfoShare$	The share of stock return variance that is attributable to trading on private firm-specific information.
Public firm-specific information share	$PublicInfoShare$	The share of stock return variance that is attributable to public firm-specific information.
Firm-specific information share	$FirmInfoShare$	The share of stock return variance that is attributable to firm-specific information (sum of $PrivateInfoShare$ and $PublicInfoShare$).
Discount rate information share	$DRShare$	The share of stock return variance that is attributable to discount rate information.
Cash flow information share	$CFShare$	The share of stock return variance that is attributable to cash flow information.
Market-wide discount rate information share	$MktInfoShare(DR)$	The share of stock return variance that is attributable to market-wide discount rate information.
Market-wide cash flow information share	$MktInfoShare(CF)$	The share of stock return variance that is attributable to market-wide cash flow information.
Private firm-specific discount rate information share	$PrivateInfoShare(DR)$	The share of stock return variance that is attributable to trading on private firm-specific discount rate information.
Private firm-specific cash flow information share	$PrivateInfoShare(CF)$	The share of stock return variance that is attributable to trading on private firm-specific cash flow information.
Public firm-specific discount rate information share	$PublicInfoShare(DR)$	The share of stock return variance that is attributable to public firm-specific discount rate information.
Public firm-specific cash flow information share	$PublicInfoShare(CF)$	The share of stock return variance that is attributable to public firm-specific cash flow information.

Appendix C: Effect of including the covariance between noise and information

In computing the variance shares in Equation (10), we ignore the covariance between information (innovations in the efficient price) and noise (changes in the pricing error). Here we show that accounting for this covariance has little effect on our estimates of the variance shares.

One way to account for the covariance term is to distribute it between the information components of variance and the noise component of variance in the same proportions as the variances of these components and then recompute the variance shares from the covariance-adjusted components using the total return variance as the normalizing variable. In this approach, we allocate a fraction α of $2cov(w_t, \Delta s_t)$ to the information variance and a fraction $(1-\alpha)$ to the noise variance, where $\alpha = \frac{\sigma_w^2}{\sigma_w^2+\sigma_s^2}$. Consequently the information and noise shares of variance become:

$$\begin{aligned}
InfoShare &= \left(\sigma_w^2 + \frac{\sigma_w^2}{\sigma_w^2+\sigma_s^2} 2cov(w_t, \Delta s_t)\right) / \sigma_r^2 \\
&= \sigma_w^2 \left(1 + \frac{2cov(w_t, \Delta s_t)}{\sigma_w^2+\sigma_s^2}\right) / \sigma_r^2 = \sigma_w^2 \left(\frac{\sigma_w^2 + 2cov(w_t, \Delta s_t) + \sigma_s^2}{\sigma_w^2+\sigma_s^2}\right) / \sigma_r^2 \\
&= \sigma_w^2 \left(\frac{\sigma_r^2}{\sigma_w^2+\sigma_s^2}\right) / \sigma_r^2 = \sigma_w^2 / (\sigma_w^2+\sigma_s^2) \\
NoiseShare &= \left(\sigma_s^2 + \frac{\sigma_s^2}{\sigma_w^2+\sigma_s^2} 2cov(w_t, \Delta s_t)\right) / \sigma_r^2 \\
&= \sigma_s^2 \left(1 + \frac{2cov(w_t, \Delta s_t)}{\sigma_w^2+\sigma_s^2}\right) / \sigma_r^2 = \sigma_s^2 \left(\frac{\sigma_r^2}{\sigma_w^2+\sigma_s^2}\right) / \sigma_r^2 \\
&= \sigma_s^2 / (\sigma_w^2+\sigma_s^2)
\end{aligned} \tag{C.1}$$

Equation (C.1) shows that after the distribution of the covariance term into information and noise components, we have exactly the same variance shares as in the baseline (Equation (10)).

An alternative way of distributing the covariance term is to add it entirely to either the information component of variance or to the noise component of variance thereby producing upper and lower bounds on the variance shares. Applying this approach we find that the upper and lower bounds are extremely narrow (e.g., the noise share has a lower bound of 25.73% and an upper bound of 27.86%, whereas the information share has a lower bound of 72.14% and an upper bound of 74.27%). Therefore, ignoring the covariance term in the baseline variance shares has little effect on the results.

References

Asparouhova, E., Bessembinder, H., Kalcheva, I., 2010. Liquidity biases in asset pricing tests. Journal of Financial Economics 96, 215–237.

Asparouhova, E., Bessembinder, H., Kalcheva, I., 2013. Noisy prices and inference regarding returns. Journal of Finance 68, 665–714.

Avramov, D., Chordia, T., Goyal, A., 2006. Liquidity and autocorrelations in individual stock returns. Journal of Finance 61, 2365–2394.

Bai, J., Philippon, T., Savov A., 2016. Have financial markets become more informative? Journal of Financial Economics 122, 625–654.

Baldauf, M., Mollner, J., 2018. High-frequency trading and market performance. Working paper.

Blume, M.E., Stambaugh, R.F., 1983. Biases in computed returns: An application to the size effect. Journal of Financial Economics 12, 387–404.

Black, F., 1986. Noise. Journal of Finance 41, 528–543.

Brogaard, J., Shi, W., Wei, K., You, H., 2018. Do analysts improve investment efficiency? Working paper.

Brogaard, J., Hendersott, T., Riordan, R., 2014. High-frequency trading and price discovery. Review of Financial Studies 27, 2267-2306.

Campbell, J.Y., 1991. A variance decomposition for stock returns. Economic Journal 101, 157–179.

Campbell, J.Y., Ammer, J., 1993. What moves the stock and bond markets? A variance decomposition for long-term asset returns. Journal of Finance 48, 3–37.

Campbell, J.Y., Lettau, M., Malkiel, B.G., Xu, Y., 2001. Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. Journal of Finance 56, 1–43.

Campbell, J.Y., Shiller, R.J., 1988a. The dividend-price ratio and expectations of future dividends and discount factors. Review of Financial Studies 1, 195–228.

Campbell, J.Y., Shiller, R.J., 1988b. Stock prices, earnings, and expected dividends. Journal of Finance 43, 661–676.

Chen, L., Da, Z., Zhao, X., 2013. What drives stock price movements?. Review of Financial Studies 26, 841–876.

Christie, W., Schultz, P., 1994. Why do NASDAQ market makers avoid odd-eighth quotes? Journal of Finance 49, 1813–1840.

Chordia, T., Roll, R., Subrahmanyam, A., 2008. Liquidity and market efficiency. Journal of Financial Economics 87, 249–268.

Chordia, T., Roll, R., Subrahmanyam, A., 2011. Recent trends in trading activity and market quality. Journal of Financial Economics 101, 243–263.

Comerton-Forde, C., Putnins, T., 2015. Dark trading and price discovery. Journal of Financial Economics 118, 70–92.

Cong, L.W., Xu, D.X., 2017. Rise of factor investing: Asset prices, informational efficiency, and security design, Working paper.

Corwin, S.A., Schultz, P., 2012. A simple way to estimate bid-ask spreads from daily high and low prices. Journal of Finance 67, 719–759.

Dasgupta, S., Gan, J., Gao, N., 2010. Transparency, price informativeness, and stock return synchronicity: Theory and evidence. Journal of Financial and Quantitative Analysis 45, 1189–1220.

DeLisle, J., French, D.W., Schutte, M.G., 2015. Co-movement, passive investing and price informativeness. Working paper.

Durnev, A., Li, K., Morck, R., Yeung, B., 2004. Capital markets and capital allocation: Implications for economies in transition. Economics of Transition 12, 593–634.

French, K., Roll, R., 1986. Stock return variances: The arrival of information and the reaction of traders. Journal of Financial Economics 17, 5–26.

Gassen, J., Skaife, H.A., Veenman, D., 2015. Illiquidity and R^2. Working paper.

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

Griffin, J., Kelly, P., Nardari, F., 2010. Do market efficiency measures yield correct inferences? A comparison of developed and emerging markets. Review of Financial Studies 23, 3225–3277.

Hasbrouck, J., 1988. Trades, quotes, inventories and information, Journal of Financial Economics 22, 229–252.

Hasbrouck, J., 1991a. Measuring the information content of stock trades. Journal of Finance 46, 179–207.

Hasbrouck, J., 1991b. The summary informativeness of stock trades: An econometric analysis. Review of Financial Studies 4, 571–595.

Hasbrouck, J., 1993. Assessing the quality of a security market: A new approach to transaction-cost measurement. Review of Financial Studies 6, 191–212.

Hendershott, T., Li, S., Menkveld, A.J., Seasholes, M.S., 2011. Risk sharing, costly participation, and monthly returns. Working paper.

Hendershott, T., Menkveld, A., 2014. Price pressures. Journal of Financial Economics 114, 405–423.

Ho, T., Stoll, H.R., 1981. Optimal dealer pricing under transaction and return uncertainty. Journal of Financial Economics 9, 47–73.

Hong, H., Kacperczyk. M., 2010. Competition and bias. Quarterly Journal of Economics 125, 1683–1725.

Hou, K., Moskowitz, T.J., 2005. Market frictions, price delay, and the cross-section of expected returns, Review of Financial Studies 18, 981–1020.

Jegadeesh, N., 1990. Evidence of predictable behavior of security returns. Journal of Finance 45, 881–898.

Jin, L., Myers, S.C., 2006. R^2 around the world: New theory and new tests. Journal of Financial Economics 79, 257–292.

Kelly, P.J., 2014. Information efficiency and firm-specific return variation. Quarterly Journal of Finance 4, 1450018.

Kelly, B., Ljungqvist, A., 2012. Testing asymmetric-information asset pricing models. Review of Financial Studies 25, 1366–1413.

Kyle, A., 1985. Continuous auctions and insider trading. Econometrica 53, 1315–1335.

Lehmann, B., 1990. Fads, martingales, and market efficiency. Quarterly Journal of Economics 105, 1–28.

Nagel, S., 2012. Evaporating liquidity. Review of Financial Studies 25, 2005–2039.

Morck, R., Yeung, B., Yu, W., 2000. The information content of stock markets: Why do emerging markets have synchronous stock price movements?. Journal of Financial Economics 58, 215–260.

Morck, R., Yeung, B., Yu., W., 2013. R^2 and the economy. Annual Review of Financial Economics 5, 143–166.

Pastor, L., Stambaugh, R.F., 2003. Liquidity risk and expected stock returns. Journal of Political Economy 111, 642–685.

Roll, R., 1984. A simple implicit measure of the effective bid-ask spread in an efficient market. Journal of Finance 39, 1127–1139.

Roll, R., 1988. R^2. Journal of Finance 43, 541–566.

Rosch, D., Subrahmanyam, A., van Dijk , M., 2017. The dynamics of market efficiency. Review of Financial Studies 30, 1151–1187.

Scholes, M., Williams, J., 1977. Estimating betas from nonsynchronous data, Journal of Financial Economics 5, 309–327.

Stoll, H.R., 1978. The supply of dealer services in securities markets. Journal of Finance 33, 1133–1151.

Subrahmanyam, A., 1991. Risk aversion, market liquidity, and price efficiency. Review of Financial Studies 4, 417–441.

Teoh, S.H., Yang, Y., Zhang, Y., 2007. R-square: Noise or firm-specific information. Working paper.

Vuolteenaho, T., 2002. What drives firm-level returns? Journal of Finance 57, 233–264.

Wurgler, J., 2000. Financial markets and the allocation of capital. Journal of Financial Economics 58, 187–214.

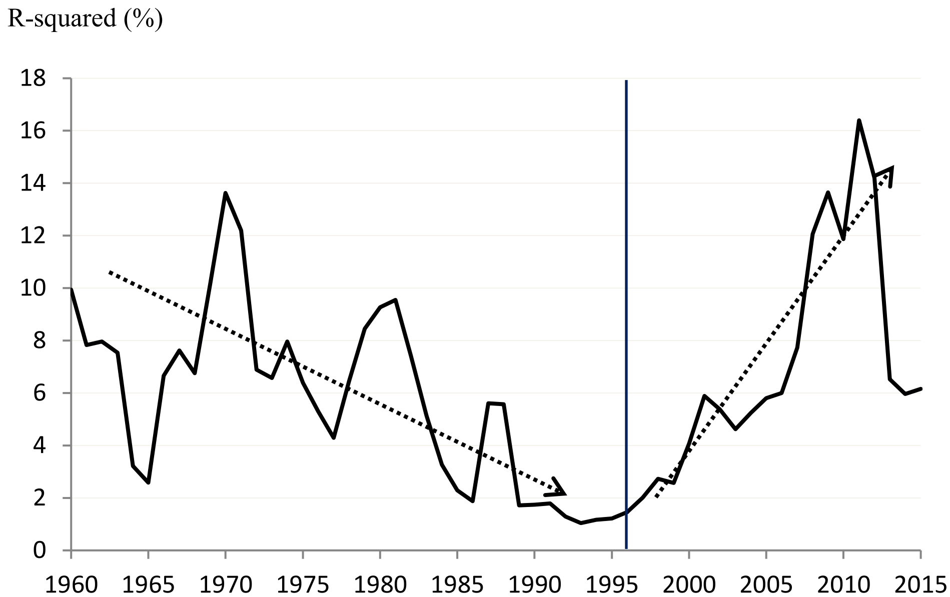


Figure 1. R^2 through time.
This figure shows the time-series trend in R^2 from 1960 to 2015. R^2 is calculated separately for each stock and each year by regressing individual daily stock returns on daily market returns, and then averaging across stocks. The sample consists of stocks listed on NYSE, AMEX, and NASDAQ.

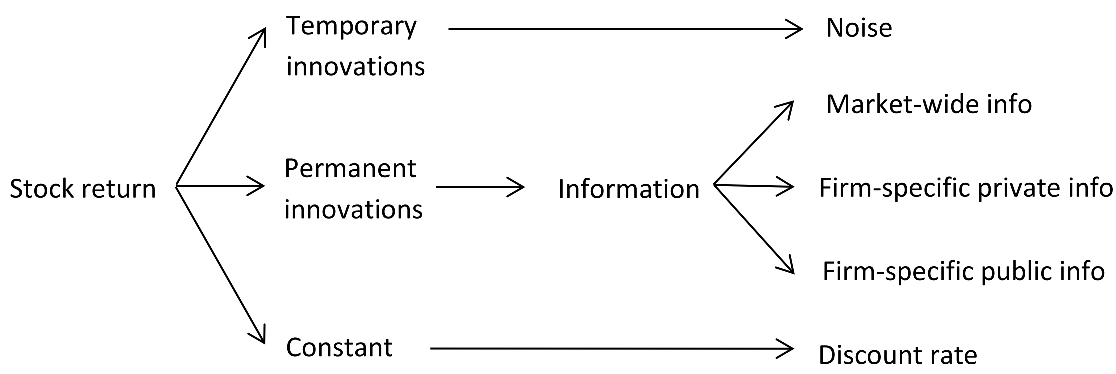


Figure 2. Stock return components in the baseline model.
In the baseline model stock returns are decomposed into temporary innovations (noise), three types of information (permanent innovations), and a constant (discount rate). The first four of these are the variance components in the baseline model, while the fifth (the discount rate) does not contribute to variance in the baseline model.

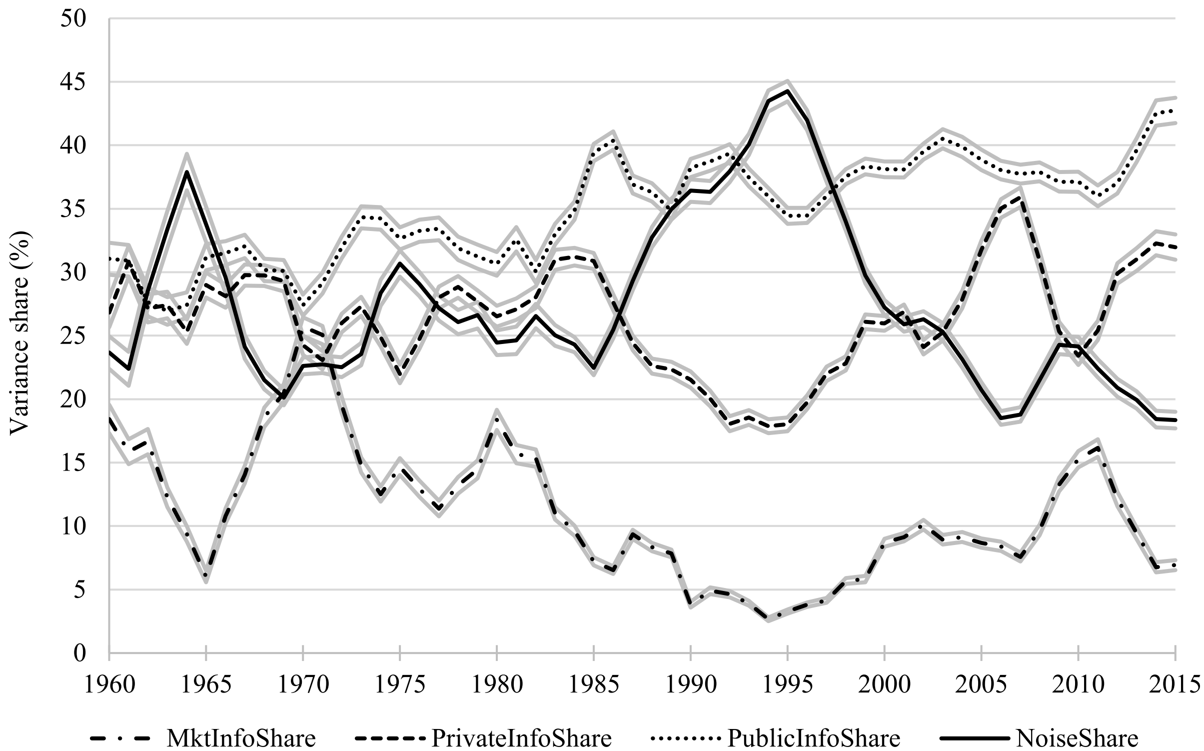


Figure 3. Stock return variance components for US stocks through time.
This figure shows the time-series trends in the percentage of stock price variance that is attributable to noise ($NoiseShare$), market-wide information ($MktInfoShare$), trading on private firm-specific information ($PrivateInfoShare$), and public firm-specific information ($PublicInfoShare$). The variance shares are calculated separately for each stock in each year based on a VAR model. We report the average variance share across stocks for each year. Light gray lines provide 99% confidence intervals. The sample consists of stocks listed on NYSE, AMEX, and NASDAQ from 1960 to 2015.

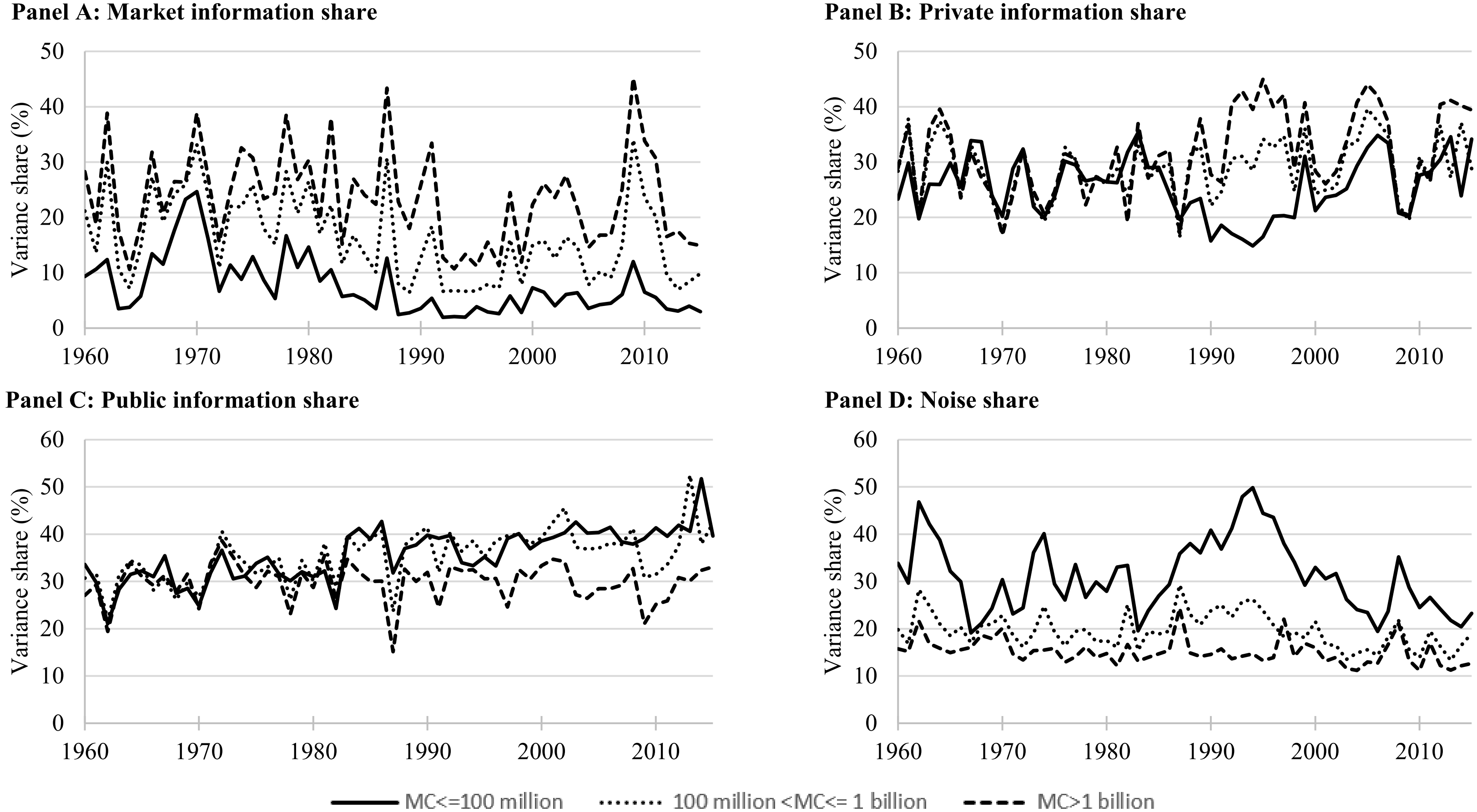


Figure 4. Variance components in size groups through time.
This figure shows the time-series trends in the percentage of stock return variance that is attributable to market-wide information (Panel A), private firm-specific information (Panel B), public firm-specific information (Panel C), and noise (Panel D) in three market capitalization groups: stocks with market capitalization less than $100 million, market capitalization between $100 million and $1 billion, and market capitalization greater than $1 billion. These breakpoint are in 2010 dollars and are adjusted for inflation forward and backward in time using the GDP price deflator. Each year stocks are assigned to one of the three groups based on their market capitalization at the start of the year. The variance component shares are calculated separately for each stock in each year and then averaged for each size group in each year. The sample consists of stocks listed on NYSE, AMEX, and NASDAQ from 1960 to 2015.

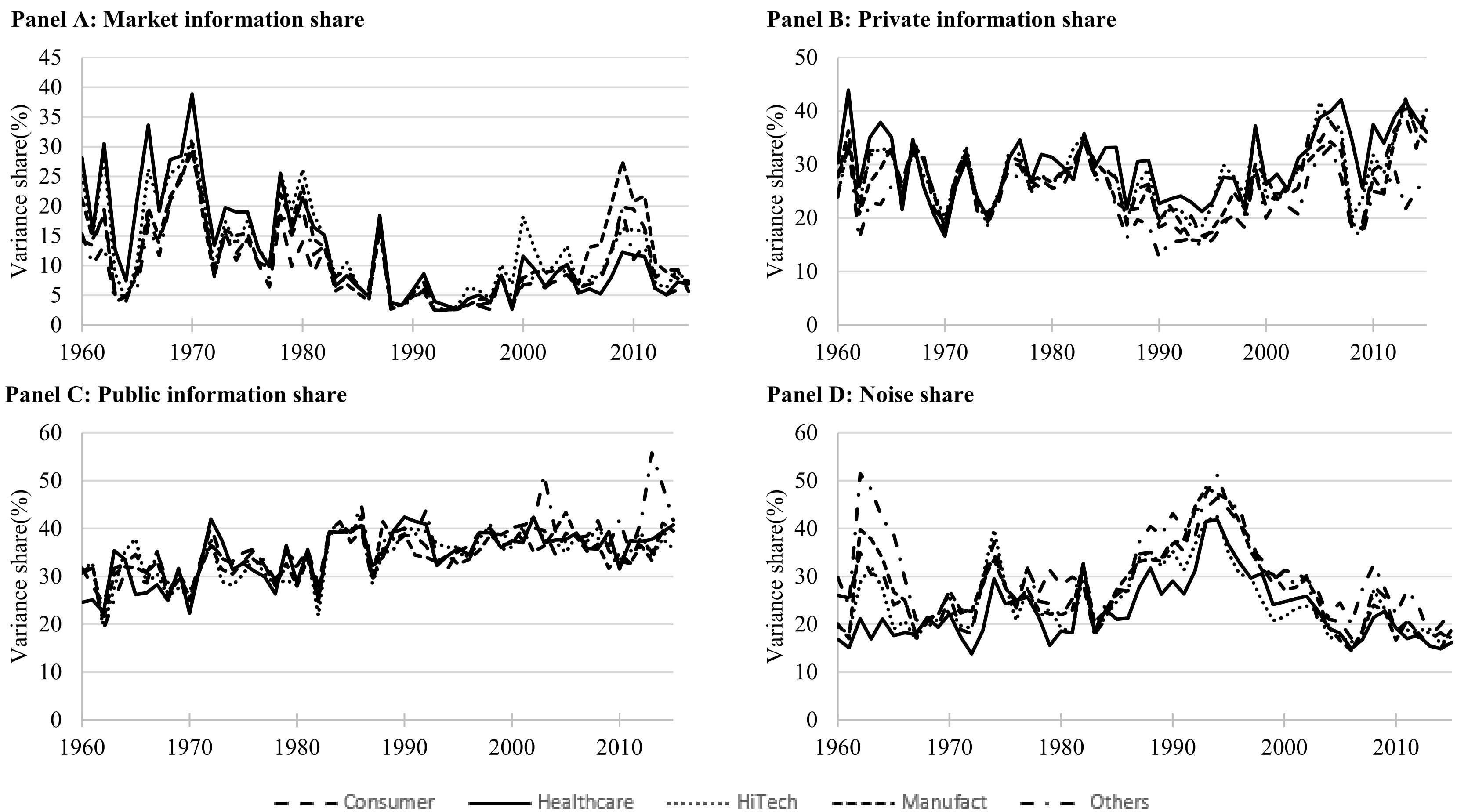
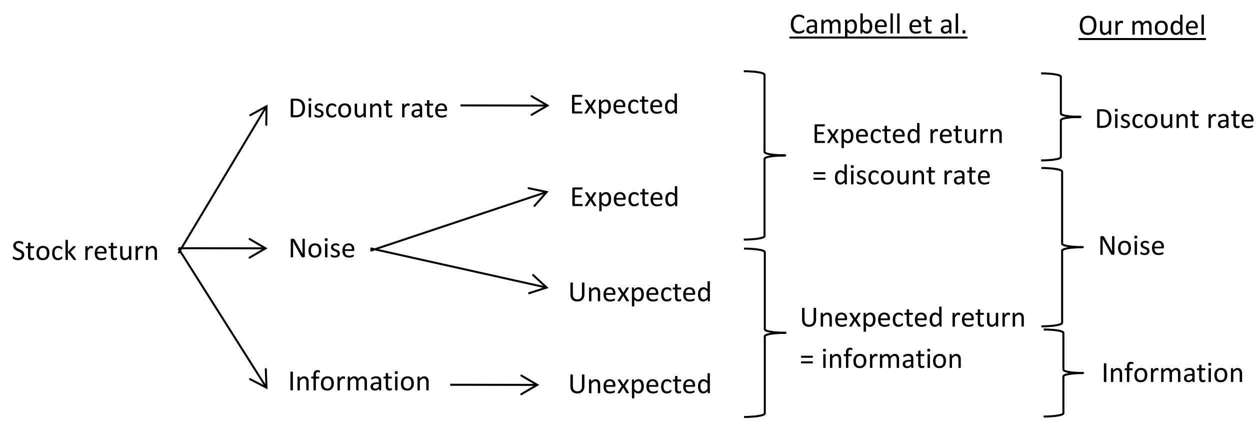


Figure 5. Variance components in major industry groups through time.
This figure shows the time-series trends in the percentage of stock return variance that is attributable to market-wide information (Panel A), private firm-specific information (Panel B), public firm-specific information (Panel C), and noise (Panel D) in five major industry groups. The *Consumer* group comprises the industries Consumer Durables, NonDurables, Wholesale, Retail, and some Services (Laundries, Repair Shops); the *Healthcare* group comprises the industries Healthcare, Medical Equipment, and Drugs; the *HiTech* group comprises the industries Business Equipment, Telephone and Television Transmission; the *Manufact* group comprises the industries Manufacturing, Energy, and Utilities; and the *Other* group comprises all other industries. The variance component shares are calculated separately for each stock in each year and then averaged for each industry group in each year. The sample consists of stocks listed on NYSE, AMEX, and NASDAQ from 1960 to 2015.

Panel A: Adjusting a standard cash flow / discount rate decomposition to account for noise



Panel B: Extended variance decomposition

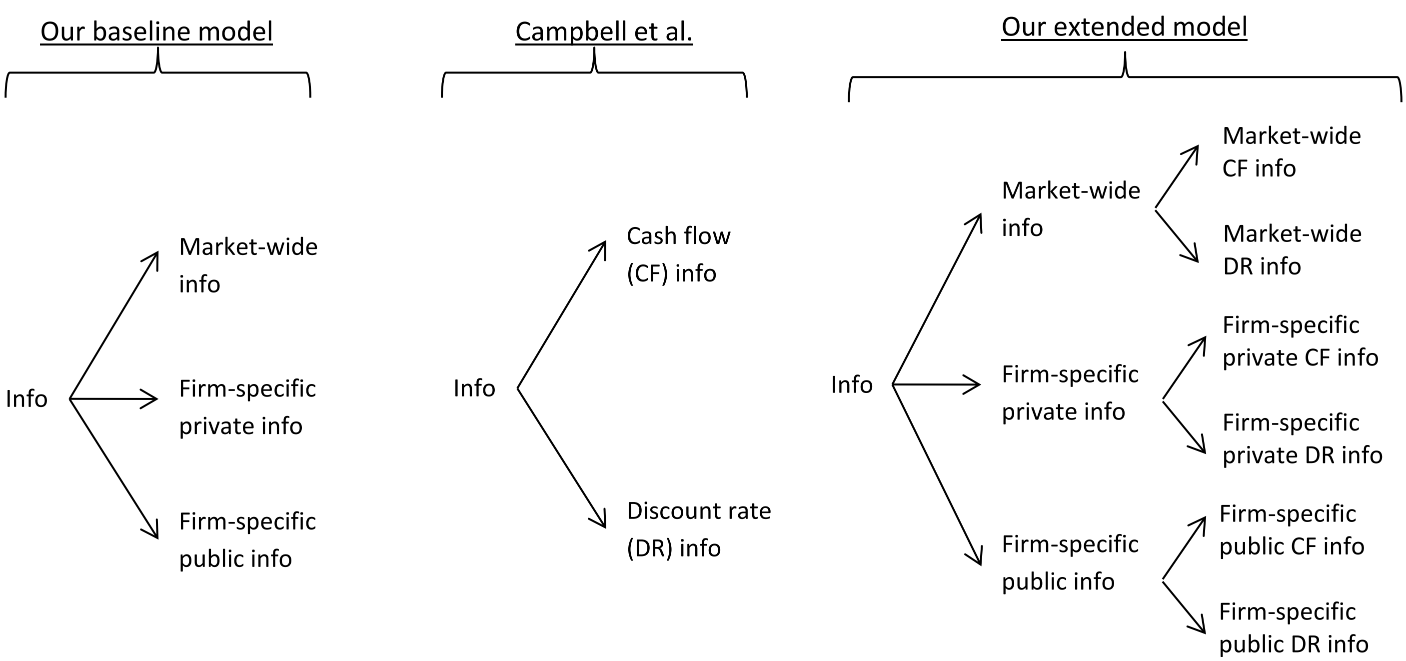
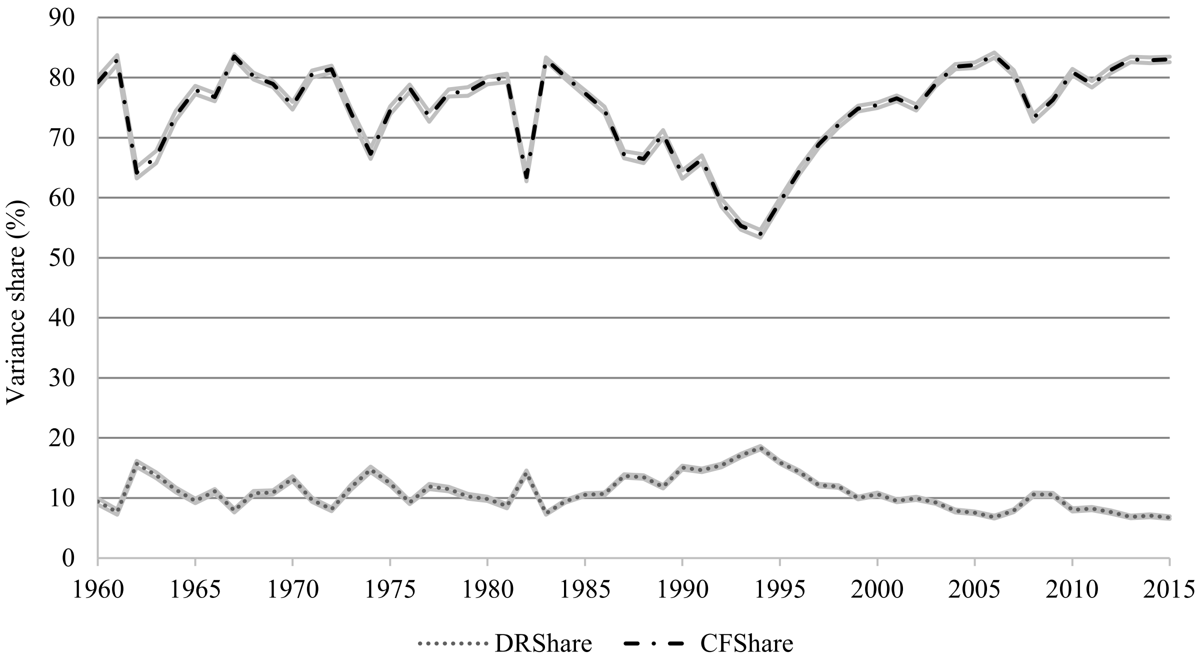


Figure 6. Extension of variance decomposition to cash flow and discount rate information.
Panel A shows how noise is dealt with in a standard cash flow / discount rate news decomposition (e.g., Campbell, 1991) and in our modified cash flow / discount rate news decomposition. In the standard decomposition, the expected changes in pricing errors contaminate the discount rate (expected return) and the unexpected changes in pricing errors contaminate the cash flow news. In our modified decomposition, noise is removed from both the discount rate and cash flow news. Panel B shows how our baseline variance decomposition is extended by splitting each of the baseline model's information components into a cash flow and discount rate part.

Panel A: Cash flow / discount rate decomposition not accounting for noise



Panel B: Cash flow / discount rate decomposition accounting for noise

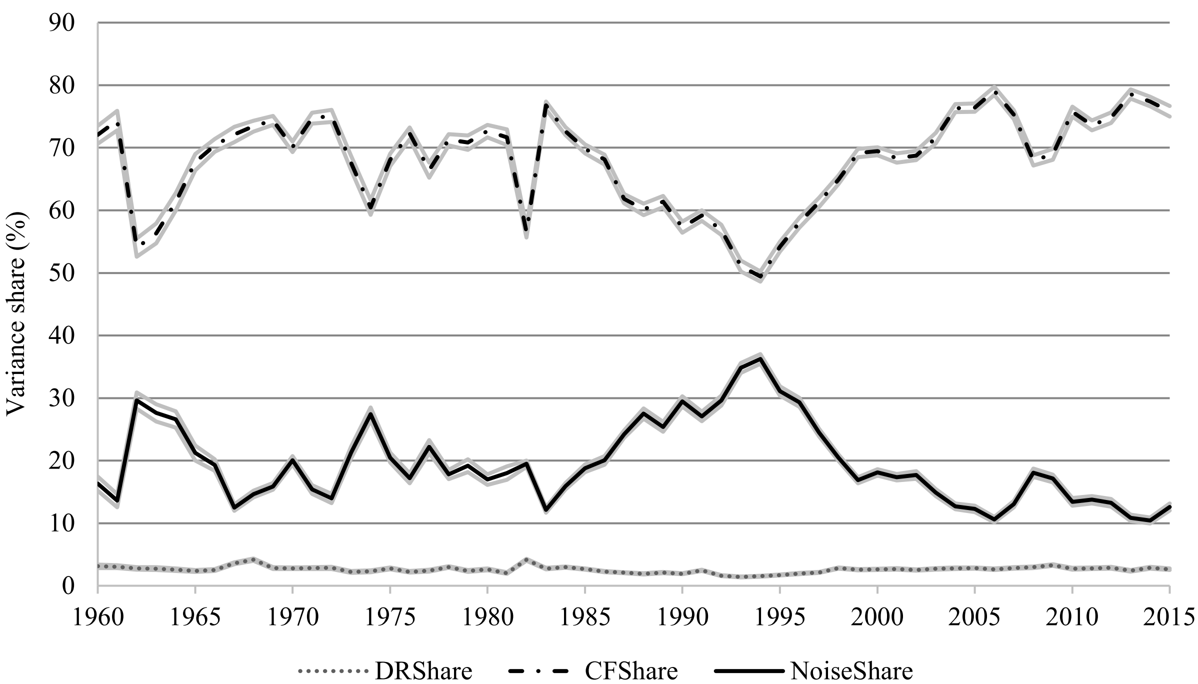


Figure 7. Cash flow news, discount rate news, and noise through time.

This figure shows the time-series trends in the percentage of stock return variance that is attributable to time-variation in the cash flow news ($CFShare$), discount rate news ($DRShare$), and noise ($NoiseShare$) from 1960 to 2015. Panel A shows the components estimated from a standard cash flow / discount rate news decomposition that does not account for noise. Panel B shows the components estimated from our modified cash flow / discount rate news decomposition that does account for noise. The variance components are calculated separately for each stock each year and then averaged across stocks each year. Light gray lines provide 99% confidence intervals. The sample consists of stocks listed on NYSE, AMEX, and NASDAQ.

Table 1. VAR coefficient estimates.

This table reports the mean coefficient estimates and the mean correlation between residuals for the baseline VAR model used to perform the variance decomposition. The VAR model is estimated separately for each stock in each year using daily observations. For the purpose of this table, each of the model coefficients is then averaged across stocks and years and reported in the table. Below each coefficient average, in parentheses, we report the percentage of negative statistically significant (at 5%) coefficients (first number in the parentheses) and the percentage of positive statistically significant (at 5%) coefficients (second number in the parentheses). The correlations column is computed similarly, but rather than reporting coefficients it reports the correlations of the residuals for pairs of variables in the VAR. The variables used in the VAR are: daily market returns in basis points ($r_{m,t}$), daily signed dollar volume in \$ thousands (x_t), and daily stock returns in basis points (r_t). The columns $l = 1$ to $l = 5$ correspond to lags of the independent variables. The sample consists of stocks listed on NYSE, AMEX, and NASDAQ from 1960 to 2015.

Dependent variable	Independent variable	$l = 1$	$l = 2$	$l = 3$	$l = 4$	$l = 5$	Correlation
Panel A: Market return equation							
$r_{m,t}$	$r_{m,t-l}$	0.104	-0.032	0.018	-0.005	-0.010	
		(2.06%, 37.93%)	(8.13%, 0.34%)	(1.69%, 3.48%)	(5.05%, 1.81%)	(2.14%, 3.09%)	
	x_{t-l}	0.069	0.024	0.033	0.081	-0.034	0.134
		(2.36%, 2.83%)	(2.51%, 2.54%)	(2.41%, 2.48%)	(2.42%, 2.63%)	(2.40%, 2.44%)	(0.57%, 42.81%)
	r_{t-l}	-0.002	-0.0001	0.0006	0.0003	-0.001	0.229
		(3.76%, 3.07%)	(3.26%, 3.19%)	(3.04%, 3.14%)	(2.90%, 3.15%)	(3.22%, 2.83%)	(0.39%, 61.85%)
Panel B: Signed dollar volume equation							
x_t	$r_{m,t-l}$	-0.650	0.013	0.612	-0.529	-0.785	0.134
		(2.27%, 10.18%)	(2.42%, 2.91%)	(1.75%, 3.50%)	(2.05%, 2.85%)	(1.97%, 2.93%)	(0.57%, 42.81%)
	x_{t-l}	0.025	-0.016	0.003	-0.008	0.004	
		(8.69%, 14.74%)	(9.21%, 5.75%)	(5.45%, 6.14%)	(5.76%, 4.52%)	(4.26%, 5.15%)	
	r_{t-l}	1.736	0.631	0.196	-0.002	0.404	0.615
		(9.55%, 8.45%)	(4.66%, 4.83%)	(3.60%, 3.82%)	(3.01%, 3.50%)	(2.79%, 3.19%)	(0.05%, 99.42%)
Panel C: Stock return equation							
r_t	$r_{m,t-l}$	0.247	0.045	0.089	0.052	0.050	0.229
		(2.10%, 21.77%)	(2.93%, 4.78%)	(1.71%, 6.02%)	(2.46%, 4.37%)	(2.36%, 4.56%)	(0.39%, 61.85%)
	x_{t-l}	0.972	-0.077	-0.059	-0.150	-0.037	0.615
		(2.56%, 11.37%)	(3.06%, 3.97%)	(2.68%, 3.10%)	(2.80%, 2.66%)	(2.65%, 2.51%)	(0.05%, 99.42%)
	r_{t-l}	-0.112	-0.060	-0.030	-0.022	-0.007	
		(31.77%, 5.47%)	(16.40%, 2.61%)	(8.88%, 2.68%)	(6.27%, 2.58%)	(4.27%, 3.07%)	

Table 2. Stock return variance components in the baseline model.

This table reports the mean variance shares (expressed as percentages of variance) for the period from 1960 to 2015. Stock return variance is decomposed into market-wide information ($MktInfoShare$), private firm-specific information ($PrivateInfoShare$), public firm-specific information ($PublicInfoShare$), and noise ($NoiseShare$). Panel A reports full sample averages. Panel B splits the sample into two sub-periods from 1960 to 1996, and from 1997 to 2015. Panel C groups stocks into quartiles by size (market capitalization) with quartiles formed separately each year. Panel D groups stocks into major industry groups: the *Consumer* group comprises the industries Consumer Durables, NonDurables, Wholesale, Retail, and some Services (Laundries, Repair Shops); the *Healthcare* group comprises the industries Healthcare, Medical Equipment, and Drugs; the *HiTech* group comprises the industries Business Equipment, Telephone and Television Transmission; the *Manufact* group comprises the industries Manufacturing, Energy, and Utilities; and the *Other* group comprises all other industries. The variance component shares are calculated separately for each stock in each year and then averaged across stocks within the corresponding quartile or group. We also report the differences in means for the post-1997 period minus the pre-1997 period (Panel B) and quartile 1 minus quartile 4 (Panel C) and report their corresponding t-statistics in parentheses. ***, **, and * indicate statistically significant differences at the 1%, 5%, and 10% levels using standard errors clustered by stock and by year. The sample consists of stocks listed on NYSE, AMEX, and NASDAQ.

	$MktInfoShare$ (%)	$PrivateInfoShare$ (%)	$PublicInfoShare$ (%)	$NoiseShare$ (%)
Panel A: Full sample				
	8.24	23.88	37.11	30.78
Panel B: Sub-periods				
1960-1996	7.00	21.89	35.63	35.47
1997-2015	9.58	26.03	38.71	25.69
Difference (Post-Pre 1997)	2.57*	4.14***	3.07***	-9.78***
	(1.73)	(2.66)	(3.37)	(-5.42)
Panel C: Quartiles by size (market capitalization)				
Q1=low	4.53	21.60	37.79	36.09
Q2	8.64	24.71	37.46	29.19
Q3	13.93	27.47	36.96	21.65
Q4=high	21.51	30.17	31.87	16.45
Difference (Q1-Q4)	-16.98***	-8.58***	5.92***	19.64***
	(-15.57)	(-5.16)	(6.16)	(12.81)
Panel D: Industry groups				
Consumer	7.69	23.24	36.49	32.59
Healthcare	7.32	28.57	37.39	26.72
HiTech	9.43	26.09	37.37	27.11
Manufact	9.15	24.29	35.02	31.55
Other	7.33	20.26	38.35	34.07

Table 3. Determinants of stock return variance components.

This table reports the results from panel regressions of stock-year observations in which the dependent variables are shares of stock return variance attributable to market-wide information ($MktInfoShare$), private firm-specific information ($PrivateInfoShare$), public firm-specific information ($PublicInfoShare$), and noise ($NoiseShare$). The explanatory variables are: D_t^{POST} is an indicator variable that takes the value of one after 1997 and zero before. $lnP_{i,t}$ is the log price and $lnMC_{i,t}$ is the log market capitalization. $D_i^{Consumer}$, $D_i^{Healthcare}$, D_i^{HiTech}, and $D_i^{Manufact}$ are indicator variables that indicate the firm's industry group (the Other Industry grouping is the omitted category). T-statistics are in parentheses using standard errors clustered by stock and by year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels. The sample consists of stocks listed on NYSE, AMEX, and NASDAQ from 1960 to 2015.

Variable	$MktInfoShare$	$PrivateInfoShare$	$PublicInfoShare$	$NoiseShare$
Intercept	-1.50**	14.48***	39.42***	47.59***
	(-2.50)	(10.13)	(32.91)	(24.35)
D_t^{POST}	-0.28	1.38	3.67***	-4.77***
	(-0.21)	(0.97)	(4.51)	(-2.83)
$lnP_{i,t}$	2.00***	-0.64	-1.09***	-0.27
	(2.90)	(-1.20)	(-2.64)	(-0.39)
$lnMC_{i,t}$	1.86***	1.93***	-0.33	-3.46***
	(4.59)	(5.66)	(-1.27)	(-9.77)
$D_i^{Consumer}$	0.53*	3.19***	-1.81***	-1.91***
	(1.92)	(7.08)	(-2.92)	(-2.71)
$D_i^{Healthcare}$	-0.48	7.38***	-1.45**	-5.44***
	(-0.91)	(8.74)	(-2.36)	(-6.62)
D_i^{HiTech}	1.27*	4.82***	-1.30**	-4.79***
	(1.85)	(6.38)	(-2.04)	(-7.92)
$D_i^{Manufact}$	1.62***	4.08***	-2.92***	-2.78***
	(4.17)	(7.51)	(-4.17)	(-3.55)
$Adjusted\ R^2$	24.1%	5.9%	1.6%	14.7%

Table 4. Effect of the tick size on noise.

This table reports the results from panel regressions of stock-year observations in which the dependent variable is the share of stock return variance attributable to noise ($NoiseShare$). Models 1-4 examine how the noise share is affected by tick size reductions from eighths of a dollar to sixteens of a dollar on June 24, 1997 using two years of data around the change (1996, 1998). $D^{POST}_{TickReduction,t}$ is an indicator variable that takes the value of one after 1997 and zero before. $Q1_i$, $Q2_i$, and $Q3_i$ are indicator variables that indicate the price quartile to which the firm belongs (the highest price quartile, $Q4_i$, is the omitted category). $lnP_{i,t}$ is the log price. Model 5 examines how the collusion by NASDAQ dealers to avoid odd-eighth quotes impacts the noise in prices, using four years before and four years during the collusion (1987 to 1994). $D^{COLLUSION}_t$ takes the value of one in the collusion period (1991-1994) and zero otherwise. $NASDAQ_i$ is an indicator variable that takes the value one for NASDAQ-listed stocks and zero otherwise. T-statistics are in parentheses using standard errors clustered by stock. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels. The sample includes all stocks listed on NYSE, AMEX, and NASDAQ.

Variable	Model 1	Model 2	Model 3	Model 4	Model 5
Intercept	19.36***	38.05***	19.67***	42.81***	20.89***
	(70.82)	(103.54)	(52.57)	(76.74)	(117.47)
$D^{POST}_{TickReduction,t}$			-0.62	-8.49**	
			(-1.54)	(-14.26)	
$D^{POST}_{TickReduction,t} \times Q1_i$			-3.43***		
			(-5.23)		
$D^{POST}_{TickReduction,t} \times Q2_i$			-4.99***		
			(-7.32)		
$D^{POST}_{TickReduction,t} \times Q3_i$			-3.94***		
			(-5.44)		
$Q1_i$	14.30***		16.02***		
	(34.65)		(27.26)		
$Q2_i$	7.56***		10.04***		
	(17.67)		(16.25)		
$Q3_i$	5.51***		7.48***		
	(13.03)		(11.92)		
$D^{POST}_{TickReduction,t} \times lnP_{i,t}$				1.72***	
				(8.38)	
$lnP_{i,t}$		-4.86***		-5.91***	
		(-36.52)		(-30.14)	
$D^{COLLUSION}_t$					-2.45***
					(-12.81)
$D^{COLLUSION}_t \times NASDAQ_i$					8.32***
					(24.89)
$NASDAQ_i$					10.02***
					(32.36)
R^2	7.5%	7.5%	8.8%	9.0%	11.7%

Table 5. Effect of analyst coverage on variance components.

This table reports the results from difference-in-difference regressions of stock-year observations in which we examine the causal effect of an exogenous drop in analyst coverage (due to brokerage mergers and closures) on variance components. The dependent variables are shares of stock return variance attributable to market-wide information ($MktInfoShare$), private firm-specific information ($PrivateInfoShare$), public firm-specific information ($PublicInfoShare$), and noise ($NoiseShare$). The independent variable of interest is $CoverageShock_{i,t}$ which is the number of broker disappearances due to mergers and closures of brokerage houses during the past two years (max of the $t-1$ and $t-2$ values). The regressions contain stock and year fixed effects. T-statistics are in parentheses using standard errors clustered by stock and year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels. The sample includes stocks listed on NYSE, AMEX, and NASDAQ from 1987 to 2011 (the period containing the analyst coverage shock events).

Variable	$MktInfoShare$	$PrivateInfoShare$	$PublicInfoShare$	$NoiseShare$
Intercept	8.02***	-3.96***	-4.77***	0.71***
	(122.82)	(-66.20)	(-110.76)	(23.13)
$CoverageShock_{i,t}$	0.43	0.57	-1.59***	0.59*
	(0.60)	(0.87)	(-3.37)	(1.80)
R^2	21.9%	6.9%	2.3%	4.9%
Fixed Effects	Stock, Year	Stock, Year	Stock, Year	Stock, Year

Table 6. Relation between variance components and other measures of information in prices.
This table reports the results from panel regressions of stock-year observations in which the dependent variables are shares of stock return variance attributable to market-wide information ($MktInfoShare$), private firm-specific information ($PrivateInfoShare$), public firm-specific information ($PublicInfoShare$), and noise ($NoiseShare$). Panel A reports the relation between the variance components and a measure of the delay with which a stock's prices respond to market-wide information ($Delay_{i,t}$). Panel B reports the relation between the variance components and a measure of the amount of information impounded in prices during the trading session ($VarianceRatio_{i,t}$ is the volatility of open to close returns divided by the volatility of overnight (close to open) returns). The sample includes stocks listed on NYSE, AMEX, and NASDAQ from 1960 to 2015 (Panel A) and 1992 to 2015 (Panel B). All regressions contain stock and year fixed effects. T-statistics are in parentheses using standard errors clustered by stock and year. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels.

Panel A: Delay metric				
Variable	$MktInfoShare$	$PrivateInfoShare$	$PublicInfoShare$	$NoiseShare$
Intercept	-2.31***	2.19***	1.60***	-1.48***
	(-18.11)	(10.12)	(7.55)	(-9.84)
$Delay_{i,t}$	-18.34***	-1.05***	4.57***	14.83***
	(-175.60)	(-6.45)	(26.08)	(87.67)
R^2	34.1%	7.3%	4.1%	11.2%
Fixed Effects	Stock, Year	Stock, Year	Stock, Year	Stock, Year

Panel B: Variance ratio				
Variable	$MktInfoShare$	$PrivateInfoShare$	$PublicInfoShare$	$NoiseShare$
Intercept	-0.62***	1.97***	1.33***	-2.68***
	(-16.60)	(49.96)	(50.29)	(-40.76)
$VarianceRatio_{i,t}$	0.44***	0.18*	0.27***	-0.89***
	(4.77)	(1.83)	(4.01)	(-5.74)
R^2	17.5%	6.2%	1.9%	6.4%
Fixed Effects	Stock, Year	Stock, Year	Stock, Year	Stock, Year

Table 7. Stock return variance components in the extended decomposition model.

This table reports mean variance shares (expressed as percentages of variance). Using an extended decomposition model, stock return variance is decomposed into market-wide information ($MktInfoShare$), private firm-specific information ($PrivateInfoShare$), public firm-specific information ($PublicInfoShare$), and noise ($NoiseShare$). The three information components are further decomposed into discount rate (DR) and cash flow (CF) related components. Panel A reports full sample averages. Panel B splits the sample into two sub-periods from 1960 to 1996, and from 1997 to 2015. Panels C and D group stocks into quartiles by price and size (market capitalization), respectively, with quartiles formed separately each year. Panel E groups stocks into major industry groups: the *Consumer* group comprises the industries Consumer Durables, NonDurables, Wholesale, Retail, and some Services (Laundries, Repair Shops); the *Healthcare* group comprises the industries Healthcare, Medical Equipment, and Drugs; the *Manufact* group comprises the industries Manufacturing, Energy, and Utilities; the *HiTech* group comprises the industries Business Equipment, Telephone and Television Transmission; and the *Other* group comprises all other industries. The variance components are calculated separately for each stock in each year and then averaged across stocks within the corresponding quartile or group. We also report the differences in means for the post-1997 period minus the pre-1997 period (Panel B) and quartile 1 minus quartile 4 (Panels C and D) and report their corresponding t-statistics in parentheses. ***, **, and * indicate statistically significant differences at the 1%, 5%, and 10% levels using standard errors clustered by stock and by year. The sample consists of stocks listed on NYSE, AMEX, and NASDAQ from 1960 to 2015.

	$MktInfoShare$ (%)		$PrivateInfoShare$ (%)		$PublicInfoShare$ (%)		$NoiseShare$ (%)
	DR	CF	DR	CF	DR	CF	
Panel A: Full sample							
	0.43	7.61	0.66	22.60	1.35	34.79	22.07
Panel B: Sub-periods							
1960-1996	0.34	6.45	0.58	20.71	1.24	33.36	26.61
1997-2015	0.52	8.85	0.75	24.63	1.48	36.34	17.16
Difference	0.18**	2.40*	0.17***	3.92***	0.24***	2.99***	-9.45***
(Post-Pre 1997)	(2.45)	(1.70)	(2.78)	(2.66)	(3.49)	(3.32)	(-6.06)
Panel C: Quartiles by Price							
Q1=low	0.30	4.56	0.64	21.50	1.36	35.36	24.66
Q2	0.47	9.00	0.67	22.61	1.41	35.30	21.27
Q3	0.69	13.75	0.73	25.23	1.29	33.71	16.25
Q4=high	0.90	19.12	0.80	27.45	1.19	30.43	12.14
Difference	-0.60***	-14.56***	-0.17***	-5.95***	0.17**	4.94***	12.52***
(Q1-Q4)	(-4.92)	(-15.96)	(-3.35)	(-4.71)	(2.47)	(6.22)	(10.90)
Panel D: Quartiles by size (market capitalization)							
Q1=low	0.28	4.05	0.59	20.29	1.31	35.28	26.29
Q2	0.44	8.04	0.69	23.49	1.42	35.22	21.25
Q3	0.67	13.02	0.78	26.22	1.42	34.90	14.58
Q4=high	0.92	20.30	0.87	28.89	1.30	30.13	9.85
Difference	-0.64***	-16.25***	-0.27***	-8.61***	0.01	5.15***	16.44***
(Q1-Q4)	(-5.10)	(-15.13)	(-4.69)	(-5.49)	(0.14)	(5.42)	(11.85)
Panel E: Industry groups							
Consumer	0.40	7.08	0.66	21.98	1.31	34.14	23.77
Healthcare	0.41	6.69	0.74	27.20	1.46	35.11	18.30
HiTech	0.47	8.74	0.71	24.68	1.37	35.14	18.75
Manufact	0.45	8.49	0.70	23.00	1.32	32.86	23.04
Other	0.40	6.73	0.59	19.12	1.34	35.87	24.90