Institutional investors and the informational efficiency of prices

Ekkehart Boehmer
eboehmer@mays.tamu.edu

Eric Kelley
ekelley@mays.tamu.edu

Christo Pirinsky
cpirinsky@mays.tamu.edu

Mays Business School
Texas A&M University
College Station, TX 77843-4218

First draft: March 1, 2005
This draft: April 18, 2005

Abstract

The fraction of domestic equity held by institutional investors has quadrupled during the past four decades, and a prominent share of trading activity is due to institutions. Yet, prior research offers diverging views on how these developments affect equity markets. In particular, we know little about how institutions affect the informational efficiency of share prices, one important dimension of market quality. We study a broad cross-section of NYSE-listed stocks between 1983 and 2003, using measures of the relative informational efficiency of prices that are constructed from transactions data. We find that stocks with a higher fraction of institutional ownership are priced more efficiently, and this result is robust across a variety of specifications. Moreover, we demonstrate that increases in actual institutional trading volume are associated with greater efficiency, and this effect appears to be distinct from the one associated with cross-sectional differences in institutional holdings.

We are grateful to Kerry Back, Paul Bennett, Joachim Grammig, and Shane Johnson for their comments and thank NYSE Research for providing part of the data.

I. Introduction

Shareholdings and trading activity by institutional investors have increased dramatically over the past decades. The percentage of U.S. equities held by members of the Securities Industries Association has risen from 16% in 1965 to 61% in 2001 (Securities Industry Association Fact Book, 2002). In addition, non-retail trading amounts to 96% of total volume in 2002 (Jones and Lipson, 2004).[1] Despite the scope of institutional ownership and trading, the consequences for the quality of equity markets remain largely unexplored.

In this paper, we study how institutional holdings, quarterly changes in holdings, and daily institutional trading are related to measures of informational efficiency in a broad cross-section of NYSE-listed stocks between 1983 and 2003. We show that institutional investors increase the relative informational efficiency of prices, which is one important dimension of market quality. The prices of stocks with a higher level of institutional ownership tend to move closer to their 'fundamental values,' and their returns resemble a random walk more closely, when compared to stocks with a lower level of institutional ownership. Using proprietary data, we also find that more intense institutional trading activity improves informational efficiency, and show that this effect is distinct from the influence of holdings per se. These results have important implications for the real economy, because more informative prices facilitate better-informed financing and investment decisions.[2]

We construct measures of relative informational efficiency from intra-day transactions data over a 21-year period. We examine how closely transaction prices track fundamental values and, in particular, test whether institutional holdings and trading affect informational efficiency in the cross-section of stocks. Following Hasbrouck (1993), we estimate the dispersion of differences

[1] Institutional trading activity is generally not publicly disclosed and this study uses proprietary audit-trail data from the New York Stock Exchange to separate retail and non-retail trading.

[2] Feedback from market prices to issuers of securities is discussed as early as Schumpeter (1912) and Keynes (1936), and underlies the q-theory of Tobin (1969). There is also extensive empirical evidence on the relation between market valuations and investment. For example, Durnev, Morck, and Yeung (2004) show that capital allocation is related to firm-specific information in returns and Wurgler (2000) presents international evidence that the link between markets and real investment is stronger in countries whose stock markets impound more firm-specific information.

between trade prices and a security's efficient price. This approach uses a vector autoregression model to separate variation of the efficient price (the random walk component of price changes) from variation of a pricing error (the stationary component). Because the pricing error is not related to fundamental information about the security, its dispersion can be interpreted as the magnitude of informational inefficiency. Intuitively, when the dispersion of pricing errors is small, equity markets incorporate information efficiently, and periods of mispricing are relatively rare. As an alternative efficiency measure, we estimate the autocorrelation of quote-midpoint returns at 30-minute and 60-minute horizons to assess how closely prices follow a random walk.

We use two data sources to measure institutional holdings and trading activity, respectively. Most institutional investors must file quarterly reports containing information on securities in which they hold a long position. From these filings we construct time series of aggregate institutional holdings and their quarterly changes in each security between 1983 and 2003. To measure trading activity, we use a proprietary data set based on audit-trail data from the New York Stock Exchange. These data allow us to compute aggregate institutional buying and selling volume on a daily basis between January 2000 and December 2003. Previous research often assumes that large institutional holdings imply higher levels of trading or uses changes in quarterly holdings to approximate trading. Our data allow us to measure the institutional share of trading volume directly, and to differentiate predictions about holdings from those about trading.

Previous empirical studies provide numerous insights on the relation between institutional trading activity and stock prices (see, for example, the review in Griffin, Harris, and Topaloglu, 2003). They disagree, however, on whether prices become more informative as a result of institutional activity. On one hand, extensive evidence suggests that institutions are better informed than other market participants and have at least some ability to forecast returns.[3] Thus, stocks with high institutional holdings or with a greater share of institutional trading should have more informative prices. On the other hand, there is also evidence that institutional trading impedes price discovery. Kothare and Laux (1995) and Sias (1996) show that greater institutional holdings increase return volatility, consistent with trading decisions that are not based on superior information. Examining the direction of causality, Sias further argues that changes in institutional holdings cause changes in volatility. Campbell, Lettau, Malkiel, and Xu (2001) compare the time-series changes of market volatility to that of individual securities, and conclude that increasing institutional holdings have contributed to greater firm-specific risk.

More closely related to our study, Sias and Starks (1997) argue that correlated trading by institutional investors contributes to the positive autocorrelation in daily stock returns between 1977 and 1991. They further show that the returns on portfolios with high institutional holdings lead (Granger-cause) those of portfolios with low holdings, which is suggestive of more efficient pricing in stocks with larger institutional holdings. We extend their analysis by measuring the institutional influence on pricing efficiency in a more direct way. Specifically, we focus on the relative magnitude of short-term deviations from efficient prices.

Our approach is motivated by the analysis in Chordia, Roll, and Subrahmanyam (CRS) (2005a). They note that order imbalances (the difference between buy and sell volume) are

---

[3] See Grinblatt and Titman (1989, 1993), Daniel, Grinblatt, Titman, and Wermers (1997), Chen, Jegadeesh, and Wermers (2000), Wermers (1999, 2000), Nofsinger and Sias (1999), and Chen, Hong, and Stein (2001). Other studies provide indirect support for the hypothesis that higher institutional ownership increases the informational content of prices. Bartov, Radhakrishnan, and Krinsky (2000) document that post-earnings announcement drift is lower for stocks with more institutional ownership; Szewczyk, Tsetsekos, and Varma (1992) and Alangar, Bathala and Rao (1999) show that high institutional ownership is associated with smaller abnormal returns subsequent to equity issues or dividend changes, respectively. Sias, Starks, and Titman (2001) argue that institutional trading provides valuable information to the market. Badrinath, Kale, and Noe (1995) show that the returns on the portfolio of stocks with the highest level of institutional ownership lead the returns on portfolios of stocks with lower levels of institutional ownership. Odean (1999) shows that the stocks purchased by individuals consistently underperform the stocks they sell, suggesting that at least some retail investors have less valuable information than other investors. On the other hand, several authors suggest that institutions may not make prices more informative even if they have better information. For example, Chevalier and Ellison (1997) find that mutual funds with poor performance experience outflows – this may limit their ability to trade against the market if prices adjust too slowly to fundamental values. Brunnermeier and Nagel (2004) discuss causes and consequences of factors that prevent arbitrage in the case of hedge funds.

predictable over several days. But contemporaneous daily stock returns are not autocorrelated, suggesting that information about future order imbalances is incorporated into prices within a trading day. Their analysis reveals that much of this information is impounded within 30 minutes and they attribute this adjustment to the activities of "astute traders," who are able to move prices with their trading activity. Most important for our analysis, CRS' study provides evidence that efficiency-creating activities tend to take place within a few minutes when markets are open. We build on this insight by focusing on intra-day periods where such activities are likely to take place, and our tests are designed to capture their success.

Based on quarterly cross-sectional regressions covering a 21-year period, we show that greater institutional holdings are associated with significantly greater informational efficiency. Moreover, efficiency varies with reported changes in institutional holdings even when conditioning on beginning-of-period holdings. These findings are robust across different measures of relative efficiency and different model specifications. We show that the efficiency improvements cannot be attributed to the increased analyst coverage that is generally associated with greater institutional holdings. One channel through which institutions may make prices more informative is through trading, and we provide support for this view. Using the time series of actual daily institutional trading volume, we show that a shock to institutional trading volume causes an immediate increase in efficiency that lasts for several days. In addition, both trading volume and the level of institutional holdings are positively related to informational efficiency in daily cross-sectional regressions. Thus, the beneficial effect of greater holdings appears to be distinct from the one associated with greater trading activity, and trading activity does not appear to be the sole channel through which institutions affect efficiency.

The mechanism that translates a better information environment into more efficient prices is not our primary concern in this study, but it is relevant to the way we interpret our results. We imagine a scenario resembling the one discussed in Chordia, Roll, and Subrahmanyam (2005a). They find that order imbalances, which are highly predictable even at a daily horizon, cease to move prices within 30 minutes. They argue that this is consistent with either market makers or attentive arbitrageurs moving prices in the direction of the expected imbalance within a few minutes. For example, in the presence of an excess of buy orders over sell orders that is expected to last for a few days, market makers might move their quotes upward or arbitrageurs may submit higher-priced sell limit orders.

In the context of institutional holdings, which are observable to market participants, these astute traders could be certain institutions (such as hedge funds), or other market participants. Holden and Subrahmanyam (1992) show that greater competition among strategic informed traders leads to faster incorporation of private information. To the extent that institutional holdings are correlated with the proportion of informed traders, this argument suggests that the informational efficiency of prices increases with institutional activity in a security. Moreover, if other market participants expect institutions to be better information producers, they should find it beneficial to be more attentive about order flow in stocks with greater institutional holdings. In particular, market makers might change the way they infer information from order flow (see Glosten and Milgrom, 1985; Kyle, 1985), or the way they balance price changes with changes in the depth of their quotes (see Kavajecz and Odders-White, 2001). Other arbitrageurs might change their order-submission strategies in a way that allows them to better adapt to changing market conditions.

Several studies present evidence that institutional trading often follows positive-feedback strategies. Positive-feedback traders increase purchases in a particular security when it has recently performed well, and sell when it has performed poorly.[4] Although such trading practices can be interpreted as rational learning through prices (Grossman and Stiglitz, 1976; Hellwig, 1980), to many they raise concerns that institutional trading could be destabilizing and trigger 'informational cascades' (Banerjee, 1992; Bikhchandani, Hirshleifer, and Welch, 1992; Welch, 1992; Avery and Zemsky, 1998; Hirshleifer and Teoh, 2001). Moreover, the known presence of feedback traders

---

[4] See Grinblatt, Titman, and Wermers (1995), Nofsinger and Sias (1999), and Cai, Kaul, and Zheng (2000). Griffin, Harris, and Topaloglu (2003) show that feedback trading can also be observed on a daily level. Sias (2004) also shows that institutional demand for a security is positively correlated with their past demand.

may prompt other investors to trade in a way that moves prices further away from their efficient values (DeLong, Shleifer, Summers, and Waldman, 1990).[5]

Our general results are inconsistent with the view that institutions move prices away from fundamental values. Moreover, we provide more direct evidence on the consequences of positive-feedback trading at quarterly and daily horizons. Using quarterly data on institutional holdings, we define changes that are in the same direction as returns over the previous quarter as momentum changes, and changes that go the opposite way as contrarian changes. In the cross-section, both momentum and contrarian changes are positively related to informational efficiency. Similarly, we condition daily institutional trading activity on previous-day returns. We find little evidence that the efficiency-enhancing effect of institutional trading differs between contrarian and momentum trades. We document, however, that purchases increase efficiency more than sells. Overall, these results imply that increases in institutional holdings and trading volume improve price discovery, even when they are based on positive-feedback strategies in the aggregate.

The remainder of this study is organized as follows. In section II, we discuss our data sources and measures of informational efficiency, institutional holdings, and institutional trading. Section III explains our empirical design. Section IV contains results using quarterly information about institutional holdings, and Section V analyzes the role of daily institutional trading for informational efficiency of prices. The final section concludes the paper.

II. Data sources and methodology

We use intra-day trade and quote data to compute alternative measures of market efficiency (described below). For securities listed on the NYSE, these data are available from ISSM between 1983 and 1992 and from the New York Stock Exchange's Trade and Quote (TAQ) database between 1993 and 2003. We match all TAQ/ISSM securities to those on CRSP on a monthly basis and, individually for each month, select all NYSE-listed domestic common stocks as our initial

---

[5] Lakonishok, Shleifer, and Vishny (1992) show that institutional trades tend to be in the same direction. Such herding raises questions similar to those we discuss in the context of feedback trading, but we do not address its consequences in this paper.

sample. Next, we obtain all primary market prices and quotes from TAQ/ISSM that satisfy standard criteria.[6] For each stock, we aggregate all trades during the same second that execute at the same price and retain only the last quote for every second if multiple quotes were issued. Between 1983 and 1998, we assume that trades are reported 5 seconds late and adjust time stamps accordingly. From January 1, 1999, we assume no reporting delay and make no time adjustment. Finally, we require that each security has at least 200 transactions per month. This leaves an average cross-sectional sample of 1,143 securities per month. The cross-section increases over time: The mean number of securities increases from 908 during the first half of the sample period, 1983-1993, to 1,402 in the second half, 1994-2003 (see Panel A in Table 1).

For each security in our sample, we compute several variables that we use to control for security-specific characteristics or market conditions. From CRSP, we compute market capitalization, consolidated trading volume, and daily closing prices. From TAQ/ISSM, we compute trade-weighted relative effective spreads, volume-weighted average prices, and the price range on a daily basis. Effective spreads are computed as twice the absolute difference between the execution price and the quote midpoint prevailing when the trade was reported (or 5 seconds earlier during the 1983-1998). The result is then standardized by the prevailing quote midpoint. The daily price range is standardized by the closing price.

Panel B in Table 1 shows descriptive statistics on each of these variables, computed as time-series averages of quarterly cross-sectional means and standard deviations. RES, the relative effective spread, decreases from 81bp during the first half of the sample to 51bp during the second half. The last two columns show that the cross-sectional dispersion of RES also decreases over time. Average trading volume, QVOL, more than triples between the two periods, and average

---

[6] We only use trades and quotes during regular market hours. For trades, we require that TAQ's CORR field is equal to zero, and the COND field is either blank or equal to *, B, E, J, or K. For ISSM, we require that the COND field is blank or equal to *, F, J, K, S, or T. We eliminate trades with non-positive prices or sizes. We also exclude a trade if its price is greater (less) than 150% (50%) of the price of the previous trade. We include only quotes that have positive depth (this filter does not apply to 1986 data, where this field is not filled) for which TAQ's MODE field is equal to 1, 2, 3, 6, 10, or 12, or for which ISSM's MODE field is equal to A, B, C, H, O, or R. We exclude quotes with non-positive ask or bid prices, or where the bid price is higher than the ask price. We require that the difference between bid and ask be smaller than 25% of the quote midpoint. We also eliminate a quote if the ask is greater than 150% of the bid.

market value increases by about 125%. In both cases, the cross-sectional dispersion increases as well over time. Share price fluctuates around $30 over the entire period.

II.1 Measuring institutional holdings and changes in holdings

Data on institutional holdings and changes in holdings originate from the 13F filings in the CDA Spectrum database. Under the 1978 amendment to the Securities and Exchange Act of 1934, all institutional investors managing a portfolio with an investment value of $100 million or more are required to file quarterly 13F reports with the SEC that list their (long) equity positions greater than 10,000 shares or $200,000 in market value as of the last date of each quarter and the corresponding change in this position since the last filing. The reported holdings represent the aggregate holdings of each reporting institution's investment subsidiaries. All of our measures based on these data are aggregated across all reporting institutions and standardized by the number of shares outstanding at the end of the quarter as reported by CRSP.[7]

Panel C in Table 1 displays summary statistics on quarterly aggregate institutional holdings, TOT, and reported changes in their holdings, TOTChg. Both measures are standardized by the number of shares outstanding. The mean holding is 49%, but it increases markedly from 43% during the first half to 55% during the second half of our sample period. This increase is accompanied by a slight increase in cross-sectional dispersion, from 18% to 21%. The mean of TOTChg is 0.64% of shares outstanding, but increases from 0.57% during the first half to 0.70% in the second half of the sample period. Similarly, its cross-sectional dispersion also increases over time.

---

[7] Because institutions only file a 13F statement if they have a sufficiently large long position, these data may contain missing values in the reported changes. This occurs if an institution fully liquidates its holdings in a stock; then it does not need to file in the next quarter, and no change in its position will be reported. In these cases we infer the change as the difference between holdings. Otherwise we use the reported changes, because they incorporate changes in the number of shares outstanding and are thus more accurate than the difference in reported holdings. We have also cross-checked adjustments for stock splits with CRSP and used the CRSP value when the 13F data were different.

II.2 Measuring daily institutional trading activity

Neither the level of holdings nor changes in holdings are necessarily good proxies for institutional trading activity. First, institutions report the net change in their position, but may have turned it over several times more. Moreover, there may be net purchases by one subsidiary and net sales of similar magnitude by another; in this case, the 13F filing would show no change. While some larger institutions operate internal markets, we generally do not observe these. Both arguments suggest that reported changes represent a lower bound on true trading activity. Second, some institutions (hedge funds, for example) may hold substantial short positions. Because short positions need not be reported, the trading volume associated with getting into and out of short positions will not be revealed by changes reported on 13F filings.[8]

We use proprietary data from the New York Stock Exchange that allows us to infer daily institutional trading volume for the period from January 2000 to December 2003.[9] The data are based on the NYSE's Consolidated Audit Trail Data (CAUD), which contains information on nearly all trades that are executed at the NYSE. CAUD is the result of matching trade reports to the underlying order data and shows for each trade the individual buy and sell orders (or market maker interest) that were executed against each other in the trade. Each of these components is identified by an "account type" variable, which provides some information on trader identity (described below). We have separately aggregated buy and sell volume for each day and security for certain combinations of account types, using the number of trades, share volume, and dollar volume. We exclude trades that are cancelled or later corrected, trades with special settlement conditions, and trades outside of regular market hours.

The account type classification is complete, because providing this information is mandatory for brokers (although it is not audited by the NYSE on a regular basis). Unfortunately,

---

[8] Two studies construct proxies for institutional trading activity from transactions data. Lee and Radhakrishna (2000) use the TORQ dataset, which provides some information about trader identities, to examine various trade-size filters based on the number of shares and dollar value of individual trade. Campbell, Ramadorai and Vuolteenaho (2004) regress 13F holdings changes on trade information from TAQ. Using estimated coefficients from these regressions, they construct daily proxies for institutional trading.

[9] Kaniel, Saar, and Titman (2004) use the same data to study retail trading.

it is still somewhat ambiguous with respect to an institutional trade classification, because it results from several different regulatory requirements. These include obligations to mark orders that are part of program trades, index arbitrage program trades, specialist trades, orders from other market makers in the stock, and short sells that are exempt from short sale restrictions. Each of these categories is further divided into proprietary member trades, trades by retail customers, and agency trades. To create a proxy for institutional trading, we take all trades except those marked as retail, proprietary member or other market maker, program trades, or specialist trades.[10]

Panel A in Table 2 reports descriptive statistics on measures of daily trading activity that we use in our tests below. Similar to Table 1, we report means and medians of cross-sectional daily averages. Table 2 is based on a smaller sample of 351 stocks, because we require a minimum of 100 trades per day to compute daily measures of informational efficiency (discussed below). On average, mean institutional volume is 58% of total NYSE volume, or 0.58% of shares outstanding. The remaining 42% of trading volume are due to program trading (22%),[11] specialist trading (16%) and retail trading (4%). Panel B in Table 2 shows average daily cross-sectional correlations between these variables and previous-quarter institutional holdings and changes. The correlation between holdings and the fraction of institutional volume is 0.15, and that between institutional turnover and holdings is 0.25. The correlation between holdings and changes in holdings and actual trading is even lower, 0.03 and 0.04, respectively. This suggests that neither holdings nor changes in holdings are meaningful proxies for institutional trading activity. In fact, the best proxy for institutional trading activity appears to be total trading volume (ρ=0.99).

[10] It is debatable whether or not program trading should be included in institutional volume. The NYSE loosely defines program trades as the trading of a basket of at least 15 NYSE securities valued at $1 million or more, without explicitly specifying a time frame that distinguishes individual from program trades. We exclude program trades, because we believe they tend to be special-purpose trades, and the motivation for conducting a program trade vary widely. For example, most of these trades are part of index arbitrage strategies and it is not clear that they are representative for the typical institutional investor in our sample. In contrast, other program trades may bundle uninformed trades, perhaps delegated by retail investors, where the bundling serves as a way to signal the absence of security-specific information. Empirically, all of our results remain qualitatively unchanged when we include program trades. Including program trades generally reduces coefficient standard errors, so we err on the conservative side.

[11] This percentage is about half of the program-trading activity reported regularly in the financial press. By convention, the publicly reported activity is computed as (buy volume + sell volume)/trading volume. Our percentage is computed as (buy volume + sell volume)/2*trading volume.

II.3 Measuring the informational efficiency of prices

Measuring how and when information is incorporated into prices has long been of interest to financial economists. Early studies use variance ratios (Barnea 1974; Hasbrouck and Schwartz, 1988) that compare long-term to short-term return variances. Relative to a unit of time, a random walk implies a ratio of one. This approach provides a simple test that can be computed from daily or monthly return data, but it is sensitive to the horizons chosen for comparison. As an alternative, other studies estimate price changes associated with new information using liquidity ratios that relate returns to volume (see Schreiber and Schwartz, 1985). While these measures are also easily computed, they do not differentiate between permanent (information-based) and temporary price changes. Because a price change due to information would be considered an efficient reaction to news, while a price change due to noise represents illiquidity, liquidity ratios are not useful as measures of informational efficiency.

In this paper we follow Hasbrouck (1993), who defines price discovery as changes in a security's efficient price. Because the efficient price is not observable, he applies a variance-decomposition procedure that empirically separates changes in the efficient price from price changes that are not related to new information. The underlying intuition is that information-based price changes should be permanent, while other price changes should be reversed quickly. His approach can be illustrated using a simple model of security price adjustment. Hasbrouck assumes that observed (log) transaction prices, p_t, can be decomposed into an efficient price, m_t, and a pricing error, s_t:

$$p_t = m_t + s_t. \tag{1}$$

The efficient price is the expectation of security value, conditional on all public information and the portion of private information that can be inferred from the current trade. It is assumed to follow a random walk whose innovations may depend on the information content of order flow, allowing market makers to react to private information revealed by orders from better-informed traders. In this model, *t* indexes transactions and not time. The pricing error may incorporate a variety of non-information related effects, including the non-information related portion of transaction costs, order imbalances, price discreteness, and dealer inventory effects. It is assumed to be a zero-mean covariance-stationary process.[12] Because the pricing error has a mean of zero, its standard error, σ_s, is a measure of its magnitude. It describes how closely transaction prices follow the efficient price over time, and can therefore be interpreted as an inverse measure of market efficiency.

Hasbrouck points out that the assumption that the efficient price follows a random walk may be problematic, because some evidence exists that returns do not follow a random walk (Lo and MacKinlay, 1988; Fama and French, 1988; Poterba and Summers, 1988). The pricing error in (1) only impounds short-term deviations from the efficient price. The length of the period over which deviations are measured depends on the actual lag structure chosen for estimation, but for practical purposes its length does not exceed a reasonable number of transactions. So if potential longer-term deviations from the efficient price are in fact temporary, the Hasbrouck measure will erroneously attribute them to changes of the efficient price, and therefore understate pricing errors. Two important features of our analysis mitigate these concerns. Most importantly, we are not using pricing errors to measure informational efficiency in an absolute sense. Rather, our analysis focuses on the relative efficiency of prices, and we are interested in factors that make the prices of a security more or less efficient. Moreover, most of our tests focus on the cross-section of stocks. Unless measurement errors implied by longer-term deviations from fundamentals are systematically related to institutional activity, this approach should make our inferences less sensitive to such concerns.[13]

Initially we estimate the pricing error on a monthly basis, but we use daily estimates for our analysis in section V. We use all trade observations except those where the reported price differs by more than 30% from the previous price. We consider these reports erroneous and delete them

---

[12] Hasbrouck (1993) allows the pricing error to be serially correlated and to be correlated with the random-walk innovation of the efficient price process.

[13] It is also worth pointing out that in this paper we are not primarily interested in whether additional information reaches the market in the form of private or public information. Our main measure of efficiency (based on Hasbrouck 1993) does not differentiate between the two, because both price variation due to public news and that due to private news inferred from order flow are attributed to changes in the efficient price.

from the sample. Following Hasbrouck (1993), we then we estimate the following VAR system with five lags and four equations:

$$\begin{bmatrix} r_t \\ \mathbf{x}_t \end{bmatrix} = \mathbf{A}_1 \begin{bmatrix} r_{t-1} \\ \mathbf{x}_{t-1} \end{bmatrix} + \mathbf{A}_2 \begin{bmatrix} r_{t-2} \\ \mathbf{x}_{t-2} \end{bmatrix} + \cdots + \begin{bmatrix} v_{rt} \\ \mathbf{v}_{xt} \end{bmatrix} \tag{2}$$

where r_t is the first difference of ln (price) and $\mathbf{x}_t$ is a three-by-one vector of the following trade variables: (1) a trade sign indicator, (2) signed trading volume, and (3) the signed square root of trading volume, allowing for a concave relationship between prices and the trade series. Following Hasbrouck (1993), we assume that a trade is buyer (seller) initiated if the price is above (below) the prevailing quote midpoint. Midpoint trades are not signed, but we include them in our estimation (with $\mathbf{x}$=0). The $\mathbf{A}_i$ are coefficient matrices, v_{rt} is the residual from the return equation and $\mathbf{v}_{xt}$ is a three-by-one vector of residuals from the trade equations. The residuals are assumed to be serially uncorrelated and to have a mean of zero. To omit overnight changes from the system, each process is restarted at the beginning of each trading day.

Given the pricing process in (1) and the vector moving average representation of (2), and identification restrictions based on Beveridge and Nelson (1981), the pricing error can be expressed as:

$$\begin{aligned} s_t = {} & \alpha_0 v_{r,t} + \alpha_1 v_{r,t-1} + \cdots + \beta_{10} v_{x1,t} + \beta_{11} v_{x1,t-1} + \cdots + \beta_{20} v_{x2,t} + \beta_{21} v_{x2,t-1} \\ & + \cdots + \beta_{30} v_{x3,t} + \beta_{31} v_{x3,t-1} + \cdots, \end{aligned} \tag{3}$$

where $v_{x1,t} - v_{x3,t}$ are elements of the $\mathbf{v}_{xt}$ vector. The α coefficients represent the pricing error's relationship to non-trade information, while the β coefficients represent its relationship to trade information; they are estimated using the impulse response coefficients from the return equation in the vector moving average representation of (2). Finally, we estimate the variance of the pricing error, σ_s^2, from:

$$\sigma_s^2 = \sum_{j=0}^{\infty} \begin{bmatrix} \alpha_j & \beta_j' \end{bmatrix} \mathrm{cov}(v) \begin{bmatrix} \alpha_j \\ \beta_j \end{bmatrix}. \tag{4}$$

We standardize σ_s by the standard deviation of $r_t=ln(p_t/p_{t-1})$ to control for cross-sectional differences in return variance and compute quarterly averages of these monthly ratios. We label this measure *V(s)/V(r)* and refer to it as the ‘pricing error’ in the remainder of this paper. To reduce the influence of outliers on our estimation, we eliminate all pricing errors that exceed one (less than 0.1% of the stock-month observations in our sample).

Panel D in Table 1 shows that the mean quarterly pricing error is 0.47, with a slightly larger median (based on quarterly averages of the monthly estimates). Over time, the average declines from 0.52 to 0.42. It is important that there is sizeable cross-section variation in this measure. The mean cross-sectional standard deviation is 0.12 and it increases slightly over time. Figure 1 shows the time series properties of this series in more detail. The four lines correspond to the cross-sectional mean, median, 25th percentile, and the 75th percentile. The vertical bars indicate the magnitude of one cross-sectional standard deviation in each direction. This graph allows several interesting observations about *V(s)/V(r)*. First, despite a gradual decline, we observe sizeable cross-sectional variation, and some variation over time. Second, the series appears well behaved in that the mean is close to the median, and the quartiles tend to lie within one standard deviation. Third, the series mirrors some developments in the market in a reasonable fashion. For example, the pricing error increases around the 1987 market crash and around the market closure during the week after 9/11/2001. We also observe a steep decline beginning in the first quarter of 2001 when decimalization was implemented. This event changed the minimum tick size from $0.0625 to $0.01. The finer pricing grid reduces the effects of price discreteness, which is one component of the pricing error. (The pricing error also declines around the change to a $1/16 pricing grid in 1997, but this decline is less pronounced than the one around decimalization.)

To test the effects of daily institutional trading volume on efficiency (Section V), we re-estimate *V(s)/V(r)* on a daily basis. To assure meaningful estimates, we require a minimum of 100 trades per day for each day and stock, which reduces the sample size compared to the quarterly analysis. Panel A in Table 2 shows that the average number of securities declines to 351, but the mean pricing error is 0.43, which is almost identical to the average for the 1994-2003 period reported in Table 1. Its average cross-sectional dispersion is 0.13, which is also identical to the larger sample. Panel B shows that both institutional volume and turnover are negatively related to *V(s)/V(r)*, indicating that more trading is associated with smaller pricing errors. Total turnover is also negatively correlated with the relative pricing error. This is potentially problematic, because institutional turnover and total turnover are almost perfectly correlated for this sample, which makes it difficult to distinguish the effects of institutional trading from those of trading activity in general. In the cross-sectional test below we address this issue by orthogonalizing total turnover with respect to institutional turnover.

Finally, we compute quote-midpoint return autocorrelations to obtain an alternative measure of informational efficiency. Using the last reported midpoint for each 30-minute (and, alternatively, 60-minute) interval (ignoring overnight returns), we compute monthly autocorrelation coefficient. Because we are interested in how closely the price series resembles a random walk, and not in the direction of the deviation, we use the logarithm of the absolute value of the autocorrelation coefficient in our estimations. While this measure does not distinguish between information related and unrelated price changes, it provides a useful comparison to the pricing-error dispersion based on Hasbrouck (1993).[14]

Panel D in Table 1 shows that – similar to the relative pricing error – the 30-minute and the 60-minute absolute measures, $|AR30|$ and $|AR60|$, decline during the sample period. Relative to the sample mean, however, the cross-sectional standard deviation is greater compared to the pricing error; so we expect cross-sectional tests to have more power using the autocorrelation measures. While not central to our analysis, we also report the levels of autocorrelation for comparison. At both sampling frequencies, mean autocorrelations are negative and decline in absolute value over time.

[14] Autocorrelation can be induced by inefficient pricing as well as efficient price discovery. For example, if informed traders split their orders over time, prices gradually incorporate information. This will induce positive autocorrelation even when all publicly available information is efficiently processed.

To illustrate the relationship between the relative pricing error and the autocorrelations, we report summary statistics on the frequency of trading in Panel E of Table 1. For each quarter, we compute the mean and median number of trades per half-hour. While trading intensity increases over time, the median of the cross-sectional means is 5.4 trades. *V(s)/V(r)* is measured in trade time, and our estimation considers five lags. |*AR30*| is measured over 30 minutes in clock time. Because the typical sample firm has about five trades during this period, the two measures cover, on average, about the same trading activity.

III. Empirical design

Our primary goal is to establish how institutional investors affect the informational efficiency of prices. If institutions can produce information at lower cost than individuals, we would expect a shift towards greater institutional holdings to increase net information production. This might manifest itself as an increase in the quantity, quality, and/or the timeliness of information about a security. Thus, greater institutional holdings should be associated with a better information environment in the cross-section of firms. Along the time dimension, the relationship between institutional holdings and efficiency is more ambiguous. Integrating a new security into an institution's (internal) research effort may take time, so that potential effects on a stock's information environment may not be instantaneous. Therefore, it is not clear that an increase in holdings is associated with a contemporaneous increase in efficiency or even an increase during the subsequent quarter. These arguments, and the properties of the series presented in Figure 1, suggest that it is more informative to study the relationship between institutions and efficiency in the cross-section, rather than along the time dimension. We examine the relationship between efficiency and institutional holdings on a quarterly basis over a 21-year sample period. In section V, we use institutional trading data over a four-year period to test whether institutional trading activity has incremental effects on efficiency, beyond those associated with the level of their holdings.

III.1 Econometric model

Our basic model uses quarterly (or daily) cross-sectional regressions of an efficiency measure on institutional holdings or trading:[15]

$$PE_{it} = \alpha_t + \beta_t I_{i,t-1} + \sum_{k=1}^{K} \gamma_{kt} X_{ki,t-1} + \varepsilon_{it} \qquad (5)$$

where PE_{it} is an estimate of the pricing error or the absolute value of the quote midpoint autocorrelation for firm i during quarter t, $I_{i,t-1}$ are measures of institutional activity in firm i during quarter $t-1$, and the X_k are a set of control variables. Throughout our analysis, we use lagged measures of institutional activity to reduce the effect that the dispersion of pricing errors may have on contemporaneous institutional holdings or trading.[16] Our hypothesis tests are based on the time series of these estimated coefficients, using Newey-West (1987) general method of moments standard errors to compute test statistics. This approach allows the relationship between pricing errors and institutional activity to vary over time and addresses several issues that commonly plague similar estimations. First, estimating separate regressions for each period minimizes the adverse effect of correlation across securities. Second, applying the Newey-West estimator (with four lags) to the time series of estimated coefficients allows coefficient variances to change over time. Third, it allows coefficients to be autocorrelated (over four periods).[17]

III.2 Control variables

We attempt to control for differences across firms that may be related to pricing efficiency. First, we include several standard controls that capture differences across firms. The logarithm of

[15] Note that all efficiency measures we use are inversely related to the degree of efficiency.

[16] The lagged measures of institutional activity can be interpreted as instruments for the corresponding current measures. Our results remain qualitatively unchanged using current measures.

[17] While an approach based on separate cross-sectional regressions is less powerful than a pooled estimation, it is also affected to a lesser extent by cross-sectional correlations among the regression errors. For example, institutions may decide to increase their exposure to a certain industry. In this case, several firms may experience similar shocks to their level of institutional holdings, which could lead to cross-sectional correlations in the errors. While this would not affect the consistency of OLS coefficient estimates, it would make the OLS estimator of the variance-covariance matrix inconsistent. Our approach avoids this problem by using only the intertemporal variation in coefficient estimates as a basis for hypothesis tests. Empirically, a panel estimator with fixed time effects and Newey-West standard errors yields qualitatively identical results for all regressions presented in this study.

market capitalization controls for differences in firm size. The logarithm of the average share price controls for a possible dependency of efficiency on the price level, for example through a greater relevance of price discreteness in lower-priced shares. Lagged trading volume controls for differences in trading activity. In robustness tests we also include lagged volatility measures and lagged quarterly buy-and-hold returns as controls. These tests are not reported because they do not qualitatively change our results.

Second, we include a measure of the average relative effective spread during the previous quarter. Effective spreads measure the magnitude of the total impact that an order has on price, and we include it for three reasons. Our dependent variable, the pricing error, measures the share of total price variance attributable to the transient component. This share could conceivably be related to the magnitude of the total price impact (the effective spread) and we wish to abstract from such scale effects. Alternatively, effective spreads are the best available measure of execution costs. Lower execution costs reduce the cost of arbitrage, and thus the costs associated with making prices more informative. Thus, controlling for effective spreads helps us to isolate efficiency improvements beyond those that are attributable to lower transaction costs. Third, both execution costs and pricing efficiency are dimensions of market quality. In a more general sense, we wish to isolate changes in pricing efficiency that go beyond simple execution-cost effects.[18]

Third, we include the lagged dependent variable in most specifications. This serves two purposes: The time series of pricing errors and midquote-return autocorrelations is relatively persistent, and we wish to confirm that the attendant autocorrelation does not affect our estimates. Moreover, institutions might conceivably base their investment decisions during quarter t-1 on the prevailing degree of efficiency during that quarter. Including lagged efficiency as a control variable should capture part of the variation in holdings that is caused by contemporaneous variation in efficiency. In general, however, we do not believe that reverse causality is as important

[18] Pricing error and effective spreads are related concepts. The pricing error is the share of the total price variance that is due to temporary effects. The effective spread measures the average magnitude of the combination of permanent and temporary price impacts. So changes in the proportion of variance that is due to temporary effects could conceivably affect how the effective spread is divided into its permanent and temporary components.

an issue in our analysis as in analyses that relate institutional holdings to returns or volatility. While it is conceivable that many institutions look at recent measures of return or volatility (see, for example, the discussion in Sias, Starks, and Titman, 2001), we find it harder to imagine that many institutions would condition trades on measures of price efficiency that are not disseminated on a regular basis.

IV. Cross-sectional results for quarterly institutional holdings

We begin with a univariate illustration of the relationship between the efficiency measures and institutional holdings. We divide our sample into size quintiles that are further divided into three groups based on the 30^{th} and 70^{th} percentile of institutional holdings. These independent sorts are performed at the beginning of each quarter, and Table 3 contains averages of the quarterly cross-sectional means. In each size quintile, each of the three efficiency measures declines from the low-holding to the medium-holding group and from the medium-holding to the large-holding group (except for $|AR60|$ in quintile 2). Thus, an increase in institutional holdings appears to be systematically associated with greater informational efficiency. This relationship seems to be stronger for lower levels of institutional holdings and for larger firms. This observation suggests that our approach of weighting firms equally in cross-sectional tests is a more conservative approach than weighting by size or institutional holdings. Moreover, it motivates including a size control and beginning-of-period level of holdings in the more rigorous tests below.

We also observe some regularity in the other variables. Institutions tend to hold a greater share in larger stocks, but we observe no systematic relationship between trading volume and holdings. Finally, institutional holdings in small firms tend to be greater in stocks that have lower average executions costs, but this relationship becomes less pronounced in the larger size quintiles.

IV.1 Effects of cross-sectional differences in institutional holdings

We now turn to a more rigorous analysis and use variants of model (5) to estimate cross-sectional regressions that relate measures of efficiency to measures of institutional holdings and their quarterly changes. We present means and medians of the quarterly regression coefficients, and test significance using a t-test and a Wilcoxon rank-sum test. While we focus on results using *V(s)/V(r)* and *ln*|*AR30*| as dependent variables, we have conducted sensitivity tests with alternative specifications. First, because of the limited distribution of *V(s)/V(r)*, we repeat all regressions using its logistic transform. Second, we repeated all tests with |*AR30*|, |*AR60*|, and *ln*|*AR60*|. Because the results are qualitatively identical in most cases, we only report these alternative specifications when the results differ materially.

Table 4 presents regressions of efficiency on controls and previous-quarter institutional holdings, *TOT*, standardized by shares outstanding. The first regression shows that lagged holdings have a significantly negative mean coefficient of -0.06 (median -0.06), controlling for *RES*, firm size, price, and trading volume. This implies that larger holdings reduce the pricing error and hence improve the efficiency of prices. Looking at the control variables, the pricing error decreases with volume and share price and increases with firm size.[19] Overall, the regression has reasonable explanatory power: the average adjusted R^2 is 0.23 for the first regression.[20]

The third regression is specified similarly, but adds the lagged pricing error as a dependent variable. As expected, its coefficient is significantly positive. But institutional holdings have still a significantly negative effect on the pricing error. However, the magnitude of the coefficient declines to -0.05 (median -0.04), indicating that part of the correlation between efficiency and holdings is caused by pricing-error persistence.

---

[19] We note that the positive coefficient on market value does not necessarily imply that larger firms are priced less efficiently, and therefore is not inconsistent with other studies. Because this coefficient represents the marginal effect of size beyond that of share price, it should not be interpreted as a measure of the total effect of size. In fact, our unconditional results in Table 3 and unreported regressions omitting share price confirm that pricing errors decrease with firm size.

[20] The following observation presents an instructive side note that aids in the interpretation of this basic result is. We find that quarterly holding coefficients in the *V(s)/V(r)* regression from 2001 to 2003 are about 10% larger in magnitude than those from 1983 to 2000 (not reported in the table). This is consistent with the pronounced decline in *V(s)/V(r)* from the first quarter of 2001 (see Figure 1) in the following sense. If we attribute the decline in *V(s)/V(r)* to decimalization, the finer pricing grid implies lower pricing errors. Yet, because the pricing grid is mandated by regulation, institutional holdings cannot influence efficiency via this component of the pricing error. Therefore, as this component's share of the error declines, the fraction of the error that is susceptible to the effect of institutional holdings should increase. This is consistent with the larger coefficients of institutional holdings after decimalization.

The remaining two regressions repeat this analysis using *ln|AR30|* as the dependent variable, again with and without its lagged value as a regressor. The results are very similar in that institutional holdings have a highly significant negative effect on departures from a random walk, which corroborates the results based on Hasbrouck's (1993) pricing error.[21] The main difference between the two efficiency measures is the sign of the coefficient on *RES*. It is zero or negative in the *V(s)/V(r)* models, but significantly positive in the *ln|AR|* regressions. If greater execution costs increase the costs of arbitrage, we would expect a positive relationship with pricing error. One explanation for the opposite sign in the *V(s)/V(r)* regressions is that increases in execution costs have a greater effect on total return variance, *V(r)*, than on the numerator – but it is difficult to disentangle these effects.

IV.2 Effects of cross-sectional changes in institutional holdings

As illustrated in Table 1, aggregate institutional holdings change over time. If cross-sectional variation in holdings affects pricing errors, we would expect that changes in holdings also have an impact. To estimate this relationship, we include aggregate changes in holdings as reported in the 13F filings, *TOTChg*, as a new regressor. Because we are interested in the marginal effect of changes, we also include the level of holdings, *TOT*, as a control. *TOT* is now lagged by two periods, so the coefficient on *TOTChg* captures the effect of ownership changes conditional on holdings at the beginning of the period.

Table 5 reports these estimates. Again, the coefficients on holdings are significantly negative and comparable in magnitude to those reported in Table 4, and the controls have similar coefficients as well. However, we find that the change in holdings has incremental explanatory power: an increase in holdings is associated with a significant decrease in pricing errors. The estimates using *ln|AR30|* as the dependent variable yield similar conclusions. Compared to the

---

[21] Sias and Starks (1997) find that institutional holdings are positively related to daily return autocorrelation, but this is not inconsistent with our finding of a negative relationship with the absolute value of (shorter-term) autocorrelation. In their analysis of individual securities, Sias and Starks show that stocks with low institutional holdings tend to have negative serial correlation, while stock with high ownership tend to have positive correlation. In unreported tests we observe the same pattern: stocks with low institutional holdings tend to have negative autocorrelations, while those with large holdings have positive ones, but they are smaller in absolute value than the negative ones.

pricing-error regressions, the effect of changes relative to that of holdings is slightly larger in the *ln|AR30|* models. Overall, these results corroborate our interpretation that institutional ownership is an important cross-sectional determinant of informational efficiency.

IV.3 Feedback trading

Institutional investors often follow positive-feedback strategies by purchasing securities following price increases and selling following price decreases (see Grinblatt, Titman, and Wermers, 1995; Nofsinger and Sias, 1999; Cai, Kaul, and Zheng, 2000). It is not clear whether feedback strategies are based on information about the security, so that their effect on pricing errors is an empirical question. To shed some light on this issue, we develop measures that condition aggregate changes in institutional holdings on buy-and-hold returns over the previous quarter. Specifically, we decompose quarterly changes in institutional holdings into momentum changes and contrarian changes. We define momentum changes, *TOTChgMOM*, as an increase (decrease) in holdings when returns during the previous quarter were positive (negative). Analogously, *TOTChgCont* contains contrarian changes: increases following negative returns, and decreases following positive returns. If momentum trades are unrelated to information about the stock, we expect a positive relationship between *TOTChgMOM* and pricing errors.

The first regression in Table 6 shows the estimated model for *V(s)/V(r)*. Again, the coefficients on control variables are similar to those in previous regressions, and the effect of *TOT* remains significantly negative. Both momentum and contrarian changes in holdings are negatively related to pricing errors. The contrarian effect is slightly larger than the momentum effect, but the coefficients are not significantly different from each other. The *ln|AR30|* estimates, however, show that only contrarian increases improve efficiency – the coefficient on *TOTChgMOM* is not significantly different from zero. This suggests that returns of stocks that are characterized by net momentum trading across all institutions are priced less efficiently within a 30-minute horizon. Finally, Table 6 also reports estimates for a *ln|AR60|* regression, because in this case the coefficient estimates are different than for the 30-minute intervals: The effect of momentum changes is still smaller in magnitude than that of contrarian changes, but within a 60-minute horizon momentum changes also have a significantly negative relationship to pricing errors.

Overall, these results suggest that increases in institutional holdings that are based on contrarian strategies benefit price discovery slightly more than increases based on momentum strategies, and the difference depends on how pricing errors are measured. However, contrarian changes appear to be associated with *faster* adjustments to informational efficiency.

IV.4 Institutional investors and analysts

So far, we have not investigated *how* institutional investors affect informational efficiency. In this section we examine one possible channel, increased analyst coverage, that may facilitate more efficiency. Brennan and Subrahmanyam (1995) show that the number of analysts following a stock is positively related to the number of institutions and their ownership. They further show that the number of analysts has a significantly negative effect on the price impact of trades, a proxy for the informational content of order flow. While the relationship between the magnitude of price impacts and relative informational efficiency is not clear, their results suggest that the number of analysts affects the intra-day information environment of a firm. Therefore, it is possible that the efficiency-increasing effect of institutional holdings arises simply because greater holdings are associated with greater analyst coverage. This explanation contrasts with our broader interpretation that institutions per se improve available information.

To differentiate between these alternative interpretations, we obtain the number of analysts that cover each firm from I/B/E/S and include its natural logarithm as a regressor. We exclude stocks without analyst forecasts in I/B/E/S. Table 7 reports estimates for the three measures of efficiency. We estimate both the holdings-only specification (as in Table 4) and the holdings-cum-changes specification (as in Table 5). Adding (lagged) analyst coverage has little effect on the coefficients of either *TOT* or *TOTChg*. They remain significantly negative in all models and for each efficiency measure. The magnitude of the holdings coefficient declines slightly in most models, but the magnitude of the changes coefficient increases slightly. Thus, controlling for analyst coverage has no measurable effect on the relationship between institutional investors and informational efficiency. This suggests that the increased analyst coverage that is associated with more institutional ownership is not the main channel through which institutions increase efficiency.

Moreover, our estimates also suggest that analysts have an effect on efficiency that is quite different in nature than the one originating from institutional investors. First, the number of analysts is not significantly related to *V(s)/V(r)* or to *ln*|*AR30*|. Only the median coefficient in the *ln*|*AR60*| regressions is significantly different from zero, and indicates that more analysts decrease autocorrelation, and thus increase efficiency. These results suggest that institutions have a more immediate effect on efficiency than analysts, perhaps because analysts are further away from the market. While this observation may be interesting in its own right, the key implication for our study is that the effect of institutions on informational efficiency appears to be quite distinct from the effect that analysts have.

V. Results on daily institutional trading

One mechanism that could directly translate greater institutional ownership into informationally more efficient prices is institutional trading activity. It is difficult to infer institutional trading from the institutional holdings reported in quarterly 13F filings. Their main purpose is to provide a snapshot of portfolio holdings and the underlying trading decisions are not directly observable in these data. Therefore, they are more appropriate for analyzing the relationship between institutional activity and informational efficiency in the cross section than over time. In contrast, the proprietary data discussed in section II.2 allows us to construct daily measures of institutional trading activity that lend themselves to an (exploratory) time-series analysis. Moreover, we combine data on daily trading with contemporaneous holdings information to separate the effects of holdings and trading in the cross-section. To guarantee meaningful estimates of daily pricing errors, all tests in this section are based on the smaller sample of (on average) 351 securities that have at least 100 trades per day.

V.1 Time-series analysis of institutional trading and informational efficiency

To gain a better understanding the dynamic relationship between institutional trading and pricing efficiency, we estimate a vector autoregression (VAR) model that relates daily innovations in institutional turnover to innovations in the relative pricing error. We aggregate *V(s)/V(r)* and institutional turnover (institutional trading volume standardized by the number of shares outstanding) by computing equally-weighted daily cross-sectional averages of both measures. We difference the pricing-error series and standardize both series using their time-series means and standard deviations.[22]

We estimate a structural VAR with five lags that allows efficiency on day *t* to be influenced by institutional turnover on day *t*, but not vice versa. This restriction is motivated by a simple economic argument about the determinants of institutional order placement decisions. Any trade on day *t* clearly affects *V(s)/V(r)* on that day instantaneously: each order changes the average pricing error once it is executed. Conversely, traders cannot condition order placement decisions on that day's pricing error, because it is not known (and may not be computable) until the close of trading. Thus, even if traders were to compute pricing errors in real time, it seems more plausible to assume that pricing errors are determined just after an order is executed.

The solid lines in Figure 2 present impulse response functions (IRFs) for this structural system, estimated over 1,003 trading days from January 1, 2000. These IRFs show the responses of each variable to a one-standard deviation shock to either its own innovation (Panel A) or to the innovation of the other variable (Panel B). The dotted lines show the cumulative responses. Each graph also includes indicator lines that mark the interval of two asymptotic IRF standard errors around zero. In Panel A, the response of turnover to a shock of its own innovation remains positive

[22] The daily pricing-error series appears to be I(1). It is characterized by non-decaying autocorrelations and Dickey-Fuller tests cannot reject the null of a random walk for lag lengths greater than two. In contrast, the aggregate quarterly pricing error and institutional activity variables at either frequency appear to be I(0). The non-stationarity of the daily pricing error is less of a concern for the cross-sectional tests below, but we use first differences, which are I(0), in the VAR analysis in this section. Specification tests also suggest that shorter lag length than five lead to autocorrelated errors.

and significant for about six weeks. In contrast, the response of daily changes in the pricing error to a shock in its own innovation dissipates quickly and is only significant on the first day.

More interesting are the cross-responses displayed in Panel B. The top half shows that an efficiency shock causes responses in turnover for about two weeks. This suggests that institutions do adapt their trading decisions to changes in informational efficiency, although the main reaction appears to be greater variability in turnover, rather than a directional effect. The responses are generally not significant. In contrast, a positive shock to turnover is associated with a significant improvement in efficiency, which is most pronounced on the subsequent trading day. About half of this improvement is reversed on the second day, and responses fluctuate around zero afterwards. The cumulative response reveals that changes to pricing errors remain below their permanent level for about two weeks following the shock to institutional turnover. Because these are responses of the changes in pricing errors, this implies a systematic and permanent decrease in pricing errors in response to an increase in institutional turnover. This finding helps interpret our cross-sectional results: One way by which institutions improve efficiency is through their actual trading activity.

V.2 Cross-sectional analysis

The quarterly cross-sectional analysis in section IV shows that greater institutional holdings appear to have a long-term, positive effect on informational efficiency. The VAR analysis reveals that greater institutional trading activity appears to have an immediate short-term effect that is positive as well. In this section, attempt to disentangle the two effects in daily cross-sectional regressions. We follow the same format as in section IV, except that we now use daily observations on the smaller sample of 351 stocks. We regress daily estimates of $V(s)/V(r)$ on (lagged) daily controls for relative effective spreads and volume-weighted average price, and quarterly controls for market value and the most recently reported institutional holdings. As before, we standardize institutional trading activity by shares outstanding. To control for changes in other volume, we also use the daily fraction of total volume that is due to institutions a measure as a measure of institutional activity.[23]

Panel A in Table 8 shows means and medians of the estimated daily regression coefficients. Among the controls, higher share price and larger market capitalization are associated with smaller pricing errors. In contrast to the quarterly estimates, *RES* now has a significantly positive coefficient, consistent with the hypothesis that higher execution costs inhibit arbitrage. Importantly, even at the daily frequency, greater institutional holdings improve efficiency significantly, and the coefficients have about the same magnitude as they do in the quarterly regressions.

The first regression uses institutional turnover to test how institutional trading affects pricing errors. The coefficient is significantly negative, consistent with the direction of the impulse responses in Figure 2. The second regression uses institutional volume as a fraction of total volume, and yields a similar conclusion: institutional trading activity appears to reduce pricing errors. Moreover, these results imply that there are distinct effects associated with the level of institutional ownership and with the actual institutional trading activity, and they do not appear to subsume each other.

---

[23] We omit the volume control in the daily regressions, because measuring non-institutional volume is nearly impossible. We would like to control for exogenous changes in trading volume. Unfortunately, as shown in Table 2, institutional volume is almost perfectly correlated with total volume. One shortcoming of our data on trading activity is that it contains no trade-by-trade information, so we cannot accurately identify the trade initiator. This makes the relationship between institutional and other volume hard to interpret. For example, suppose that most institutional volume comes from informed traders who have just received a private signal. Then most institutions would wish to trade in the same direction. To complete these trades, they would need to compensate other traders for providing liquidity and taking the other side. Therefore, the increase in institutional volume would be matched by an increase in other volume. Because we do not observe who initiates trades, we cannot differentiate between volume due to information, volume that is induced by greater liquidity premiums, volume arising from traders mimicking institutions, and volume due to exogenous factors. This issue becomes even more complex when institutions represent more than 50% of volume, as in our sample, so that it is not possible for all institutions to trade on the same side.

We conducted sensitivity tests that employ controls for non-institutional volume. In a first test, we include the logarithm of non-institutional volume as a regressor. Second, we construct a measure of non-institutional turnover that is not correlated with institutional turnover by regressing total turnover on institutional turnover; then we include the regression error in the turnover regressions. The regression error represents the portion of total turnover that is unrelated to institutional turnover, and is a measure of non-institutional trading activity. These additional regressions are not reported, because none of them materially affects the coefficient estimates reported in Table 8.

Finally, we condition institutional trading on trade direction and the return on the previous trading day. Similar to the quarterly analysis in section IV.3, we test whether momentum and contrarian trades affect pricing errors in the same way. In addition, we differentiate between buys and sells. Prior evidence shows that buys have larger price impacts than sells, suggesting that buys are more informative.[24] Other things equal, we would therefore expect that buys have a stronger effect on informational efficiency than sells do.

The data contain separate variables for buying and selling activity. We define momentum buys as purchases following a positive return on the previous day, and contrarian buys as purchases following a negative return on the previous day. The sell variables are defined analogously, and trading activity following zero returns is set to zero. Panel B in Table 8 presents the coefficient estimates for the two regression models that use shares outstanding and total volume, respectively, to standardize institutional trading activity. Each trading measure is negatively related to pricing errors, so that institutional trading improves efficiency regardless of trade direction and whether trades are based on a feedback strategy. Consistent with the quarterly results on changes in holdings, we find little evidence that momentum trades affect efficiency in a different way than contrarian trades, although momentum sells are not significantly related to efficiency in the turnover regression. But the estimates show that institutional buying has a stronger effect on efficiency than institutional selling. The magnitudes of the coefficients on institutional buying are between 50% and 200% larger than those on selling, and most of these differences are significant at the five-percent level. This is consistent with the buy-sell asymmetry discussed above. Overall, institutional trading activity appears to have beneficial consequences for the informational efficiency of prices, and this effect does not subsume the one associated with greater institutional holdings.

---

[24] Buy transactions tend to be more informative and costlier to complete than sell transactions. See Chan and Lakonishok (1993), Gemmil (1996), Keim and Madhavan (1997), and Kalay, Sade, and Wohl (2004) for evidence from different stock markets. Saar (2001) provides a theoretical analysis of the asymmetry between buys and sells.

VI. Conclusions

Institutional investors' share holdings and their contribution to trading volume have increased substantially over the past two decades. Yet, relatively little is known about how their presence affects the informational efficiency of prices, one important dimension of market quality. We study this issue using a broad sample of NYSE stocks between 1983 and 2003. We use all intraday transactions during this period to compute different measures of relative efficiency. We then relate these measures to institutional holdings, quarterly changes in holdings, and daily institutional trading activity.

We find that a greater share of institutional holdings is associated with greater informational efficiency of prices in the cross-section of stocks. In addition, signed quarterly changes in institutional holdings are positively related to informational efficiency. This suggests that the presence of institutional investors improves the informational environment of a firm. Because our efficiency measures concentrate on short-run deviations from efficient prices, the presence of institutional investors appears to make market participants who are close to the trading process more attentive. We provide some direct evidence that institutions themselves have a role in this process, because their trading activity is directly associated with efficiency improvements. But we also document that the magnitude of institutional holdings per se has additional influence on efficiency. We interpret this as evidence that other market participants, perhaps market makers, also become more attentive when institutional ownership increases.

While our measures of the relative efficiency of prices concentrate on the short run, we believe that the effects we document are distinct from the dimensions of market quality that are typically assessed in studies of market microstructure. In particular, we control for the magnitude of execution costs throughout our analysis, so the effect we ascribe to institutional investors goes beyond simple reductions in trading costs.

Previous studies document that institutions are often positive-feedback traders, and speculate about the consequences for the informational content of market prices. Our quarterly evidence shows that increases in institutional holdings improve efficiency even when based on positive-feedback strategies, although in a slightly different manner than changes based on contrarian strategies. At a daily horizon, where we can measure institutional trading directly, we also find little difference in the efficiency-enhancing effects of momentum and contrarian trading by institutions. These findings are consistent with the view that positive-feedback is the result of an economically meaningful strategy, such as order splitting, rather than a manifestation of irrational behavior.

Because greater institutional ownership increases analyst coverage, an alternative explanation for our results might attribute efficiency improvements to financial analysts, rather than directly to institutions. Our results are inconsistent with this view, because they remain largely unchanged when we control for analyst coverage. We show that analyst coverage has little effect on informational efficiency, and our tests suggest that their impact is quite distinct from that of institutional investors.

In a broader sense, the informational efficiency of prices is a valuable public good, because all market participants benefit from more efficient prices. Our finding that institutions increase efficiency contributes to the ongoing debate on how their increasing presence affects the quality of equity markets. It seems important for future research efforts to identify the channels through which these efficiency improvements materialize, because the precise nature of this process should have important implications for the optimal design of markets, trading protocols, and regulatory policy.

References

Alangar, S., C. Bathala, and R. Rao, 1999, The effect of institutional interest on the informational content of dividend-change announcements, Journal of Financial Research 22, 429-448.

Avery, C. and P. Zemsky, 1998, Multidimensional uncertainty and herd behavior in financial markets, American Economic Review 88, 724-748.

Badrinath, S., J. R. Kale, and T. Noe, 1995, Of shepherds, sheep, and the cross-autocorrelations in equity returns, Review of Financial Studies 8, 401-430.

Banerjee, A., 1992, A simple model of herd behavior, Quarterly Journal of Economics 107, 797–817.

Barnea, A., 1974, Performance evaluation of New York Stock Exchange specialists, Journal of Financial and Quantitative Analysis 9, 511-535.

Bartov, E., S. Radhakrishnan, and I. Krinsky, 2000, Investor sophistication and patterns in stock returns after earnings announcements, Accounting Review 75, 43-63.

Beveridge, S. and C. Nelson, 1981, A new approach to the decomposition of economic time series into permanent and transitory components with particular attention to the measurement of the 'business cycle,' Journal of Monetary Economics 7, 151-174.

Bikhchandani, S., Hirshleifer D., and I. Welch, 1992, A theory of fads, fashion, custom, and cultural change as informational cascades, Journal of Political Economy 100, 992-1026.

Brennan, M. J. and A. Subrahmanyam, 1995, Investment analysis and price formation in securities markets, Journal of Financial Economics 38, 361-382.

Brunnermeier, M. K. and S. Nagel, 2004, Hedge funds and the technology bubble, Journal of Finance 59, 2013-2040.

Cai, F., G. Kaul, and L. Zheng, 2000, Institutional trading and stock returns, working paper, University of Michigan.

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

Campbell, J. Y., T. Ramadorai, and T. O. Vuolteenaho, 2004, Caught on tape: Predicting institutional ownership with order flow, Working paper, Harvard University.

Chan, L. K. C. and J. Lakonishok, 1993, Institutional trades and intraday stock price behavior, Journal of Financial Economics 33, 173-199.

Chen, H., N. Jegadeesh, and R. Wermers, 2000, The value of active mutual fund management: An examination of the stockholdings and trades of fund manages, Journal of Financial and Quantitative Analysis 35, 343-368.

Chen, J., H. Hong, and J. Stein, 2001, Breadth of ownership and stock returns, working paper, Harvard University.

Chevalier, Judith, and Glenn Ellison, 1997, Risk taking by mutual funds as a response to incentives, Journal of Political Economy 105, 1176–1200.

Chordia, T., R. Roll, and A. Subrahmanyam, 2005a, Evidence on the speed of convergence to market efficiency, Journal of Financial Economics, forthcoming.

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers, 1997, Measuring mutual fund performance with characteristic based benchmarks, Journal of Finance 52, 1035-1058.

DeLong, J. B., A. Shleifer, L. Summers, and R. Waldmann, 1990, Positive feedback investment strategies and destabilizing rational speculation, Journal of Finance 45, 379-395.

Durnev, A., R. Morck, and B. Yeung, 2004, Value-enhancing capital budgeting and firm-specific stock return variation, Journal of Finance 59, 65-105.

Fama, E. F., and K. R. French, 1988, Permanent and temporary components of stock prices, Journal of Political Economy 96, 246-273.

Gemmil, G., 1996, Transparency and liquidity: A study of block trades on the London Stock Exchange under different publication rules, Journal of Finance 51, 1765-1790.

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

Griffin, J. M., J. H. Harris, and S. Topaloglu, 2003, The dynamics of institutional and individual trading, Journal of Finance, forthcoming.

Grinblatt, M. and S. Titman, 1989, Portfolio performance evaluation: Old issues and new insights, Review of Financial Studies 2, 393-422.

Grinblatt, M. and S. Titman, 1993, Performance measurement without benchmarks: An examination of mutual fund returns, Journal of Business 66, 47-68.

Grinblatt, M., S. Titman, and R. Wermers, 1995, Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior, American Economic Review 85, 1088-1105.

Grossman, S. and J. Stiglitz, 1976, Information and competitive price systems, American Economic Review 66, 246-253.

Hasbrouck, J. and R.A. Schwartz, 1988, An assessment of stock exchange and over-the-counter markets, Journal of Portfolio Management 14, 10-16.

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

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

Hirshleifer D. and S. Teoh, 2001, Herd behavior and cascading in capital markets: A review and synthesis, Working paper, Ohio State University.

Holden, C. W. and A. Subrahmanyam, 1992, Long-lived private information and imperfect competition, Journal of Finance 47, 247-270.

Jones, C. M. and M. L. Lipson, 2004, Are retail orders different? Working paper, Columbia University.

Kalay, A., O. Sade, and A. Wohl, 2004, Measuring stock illiquidity: An investigation of the demand and supply schedules at the TASE, Journal of Financial Economics 74, 461-486.

Kaniel, R., G. Saar, and S. Titman, 2004, Individual investor sentiment and stock returns, Working paper, Duke University.

Kavajecz, K. A. and E. R. Odders-White, 2001, An examination of changes in specialists' posted price schedules, Review of Financial Studies 14, 681-704.

Keim, D. B. and A. Madhavan, 1997, Transaction costs and investment style: An inter-exchange analysis of institutional equity trades, Journal of Financial Economics 46, 265-292.

Keynes, J. M., 1936, The general theory of employment, interest, and money, Harvest Book edition, Harcourt Brace.

Kothare, M. and P. A. Laux, 1995, Trading costs and the trading systems for Nasdaq stocks, Financial Analyst Journal 51, 42-53.

Kyle, A. S., 1985, Continuous auction and insider trading, Econometrica 53, 1315-1336.

Lakonishok, J., A. Shleifer, and R. W. Vishny, 1992, The impact of institutional trading on stock prices, Journal of Financial Economics 32, 23-43.

Lee, C. M.C. and B. Radhakrishna, 2000, Inferring investor behavior: Evidence from TORQ data, Journal of Financial Markets 3, 83-112.

Lo, A.W. and A.C. MacKinlay, 1988, Stock prices do not follow a random walk: Evidence from a simple specification test, Review of Financial Studies 1, 41-66.

Newey, W. and K. West, 1987, A simple positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix, Econometrica 55, 703-708.

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

Odean, T., 1999, Do investors trade too much? American Economic Review 89, 1279-1298.

Poterba, J. M. and L. H. Summers, 1988, Mean reversion in stock prices: Evidence and implications, Journal of Financial Economics 22, 27-60.

Saar, G., 2001, Price impact asymmetry of block trades: An institutional trading explanation, Review of Financial Studies 14, 1153-1181.

Schreiber, P. S., and R. A. Schwartz, 1986, Price discovery in securities markets, Journal of Portfolio Management 12, 43-48.

Schumpeter, J. A., 1912, Theorie der Wirtschaftlichen Entwicklung: Eine Untersuchung ueber Unternehmergewinn, Kaptial, Kredit, Zins und den Konjunkturzyklus, Duncker and Humblot, Berlin.

Sias, R. W., 1996, Volatility and the institutional investor, Financial Analysts Journal 52, 13-20.

Sias, R. W., 2004, Institutional herding, Review of Financial Studies, forthcoming.

Sias, R. W. and L. Starks, 1997, Return autocorrelation and institutional investors, Journal of Financial Economics 46, 103-131.

Sias, R. W., L. Starks, and S. Titman, 2001, The price impact of institutional trading, Working paper, University of Texas.

Szewczyk, S., G. Tsetsekos, and R. Varma, 1992, Institutional ownership and the liquidity of common stock offerings, Financial Review 27, 211-225.

Tobin, J., 1969, A general equilibrium approach to monetary theory, Journal of Money, Credit and Banking 1, 15-29.

Welch, I., 1992, Sequential sales, learning, and cascades, Journal of Finance 47, 695-732.

Wermers, R., 2000, Mutual fund performance: An empirical decomposition into stock-picking talent, style, transaction costs, and expenses, Journal of Finance 55, 1655-1695.

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

Table 1: Descriptive statistics

The sample is based on NYSE-listed securities between 1983 and 2003. We compute cross-sectional means and standard deviations for every quarter, and this table report statistics on the time series of these estimates. For example, Mean Std is the time-series average of the 84 quarterly cross-sectional standard deviations. V(s)/V(r) is the relative pricing error based on Hasbrouck (1993). AR30 is the 30-minute quote midpoint return autocorrelation, and AR60 the corresponding measure for 60-minute returns. TOT is the fraction of shares outstanding held by institutions who file quarterly 13F reports. TOTChg is the percentage of shares outstanding by which these institutions' aggregate holdings have changed during a quarter. RES is the trade-weighted relative effective spread. QVOL is share trading volume, MV is the market value of equity, and Price is the share price. NoTrd30min is the number of trades within 30 minutes.

	Mean	Median	Mean Std	Mean 1983-1993	Mean 1994-2003	Mean Std 1983-1993	Mean std 1994-2003
Panel A: Number of securities included							
N	1,143	1,110		908	1,402		
Panel B: Control variables							
RES	0.67%	0.69%	0.83%	0.81%	0.51%	0.99%	0.65%
QVOL (round lots)	217,505	138,547	426,000	104,063	342,292	139,636	741,000
MV ($1,000)	3,705,786	3,114,833	10,343,952	2,332,558	5,216,337	5,159,648	16,046,687
Price ($)	30.96	31.06	24.34	31.51	30.35	22.92	25.91
Panel C: Measures of institutional holdings and net changes in holdings, standardized by shares outstanding							
TOT	49%	49%	19%	43%	55%	18%	21%
TOTChg	0.64%	0.62%	4.98%	0.57%	0.70%	4.40%	5.62%
Panel D: Measures of market efficiency							
V(s)/V(r)	0.472	0.493	0.115	0.516	0.424	0.106	0.126
|AR30|	0.071	0.069	0.056	0.075	0.066	0.060	0.052
|AR60|	0.088	0.084	0.093	0.099	0.076	0.122	0.062
AR30	-0.019	-0.019	0.086	-0.026	-0.011	0.091	0.080
AR60	-0.009	-0.009	0.131	-0.015	-0.002	0.163	0.096
Panel E: Measures of trading frequency							
Mean NoTrd30min	11.8	5.4		4.4	20.0		

Table 2: Descriptive statistics on daily institutional trading

The sample is based on NYSE-listed securities between Jan 1, 2000 and Dec 31, 2003. We compute cross-sectional means and standard deviations for every trading day, and this table report statistics on the time series of these estimates. For example, Mean Std is the time-series average of the 1005 daily cross-sectional standard deviations. The sample contains only stocks with at least 100 trades per day. N is the average number of securites per day. V(s)/V(r) is the daily relative pricing error based on Hasbrouck (1993). TOT is the fraction of shares outstanding held by institutions who file quarterly 13F reports, and TOTChg is the reported quarterly change from these reports. Ivol is the institutional trading volume executed on the NYSE. TotVol is total volume executed on the NYSE, and Shrout is the number of shares outstanding.

Panel A: Descriptive statistics

	Mean	Median	Mean Std
N	351	351	
V(s)/V(r)	0.43	0.41	0.13
Ivol/TotVol	0.58	0.57	0.10
Ivol/Shrout (10 bp = 1 unit)	5.84	5.83	6.87

Panel B: Correlations of institutional holdings and trading

	V(s)/V(r)	TOT	TOTChg	Ivol/TotVol	Ivol/Shrout	TotVol/Shrout
V(s)/V(r)	1.000					
TOT	-0.002	1.000				
TOTChg	0.010	0.061	1.000			
Ivol/TotVol	-0.062	0.148	0.032	1.000		
Ivol/Shrout	-0.053	0.254	0.038	0.508	1.000	
TotVol/Shrout	-0.053	0.282	0.037	0.425	0.986	1.000

Table 3: Summary statistics by market value and institutional ownership

The sample is based on NYSE-listed securities between 1983 and 2003. Each quarter, firms are divided into quintiles based on the market value of equity at the beginning of the quarter. Within each size quintile, firms are divided into three groups based on institutional ownership standardized by the number of shares outstanding, also at the beginning of the quarter. The table presents means over 84 quarters of the cross-sectional means computed in each quarter. V(s)/V(r) is the relative pricing error based on Hasbrouck (1993). |AR30| is the absolute value of the 30-minute quote midpoint return autocorrelation, and |AR60| the corresponding measure for 60-minute returns. RES is the trade-weighted relative effective spread, QVOL is share trading volume, Price is the share price, TOT is the fraction of shares outstanding held by institutions who file quarterly 13F reports, and MV is the market value of

	Institutional holdings / shares outstanding		
	< 30th percentile	30th-70th percentile	>70th percentile
Size quintile 5 (largest)			
V(s)/V(r)	0.458	0.419	0.385
|AR30|	0.063	0.054	0.053
|AR60|	0.074	0.068	0.067
RES	0.003	0.003	0.003
Price	42.96	49.29	55.09
QVOL (round lots)	717,716	660,427	471,883
TOT	0.360	0.559	0.717
MV (000,000s)	16,962	13,326	7,449
Size quintile 4			
V(s)/V(r)	0.487	0.444	0.425
|AR30|	0.077	0.062	0.060
|AR60|	0.089	0.077	0.075
RES	0.005	0.004	0.004
Price	28.89	33.43	36.93
QVOL (round lots)	156,516	197,914	205,740
TOT	0.301	0.546	0.728
MV (000,000s)	1,650	1,744	1,745
Size quintile 3			
V(s)/V(r)	0.519	0.478	0.470
|AR30|	0.089	0.068	0.065
|AR60|	0.104	0.089	0.084
RES	0.007	0.006	0.006
Price	23.49	24.93	27.92
QVOL (round lots)	80,022	96,404	101,690
TOT	0.228	0.482	0.694
MV (000,000s)	667	672	681
Size quintile 2			
V(s)/V(r)	0.535	0.511	0.508
|AR30|	0.094	0.078	0.071
|AR60|	0.111	0.100	0.111
RES	0.013	0.009	0.008
Price	14.85	17.70	19.54
QVOL (round lots)	55,901	53,999	52,279
TOT	0.197	0.419	0.641
MV (000,000s)	267	276	287
Size quintile 1 (smallest)			
V(s)/V(r)	0.555	0.535	0.534
|AR30|	0.125	0.105	0.086
|AR60|	0.144	0.126	0.107
RES	0.036	0.021	0.015
Price	6.10	8.79	11.41
QVOL (round lots)	55,867	43,062	35,546
TOT	0.110	0.272	0.497
MV (000,000s)	78	96	111

Table 4: The cross-sectional effect of lagged institutional holdings on pricing errors

The sample is based on NYSE-listed securities between 1983 and 2003. We conduct quarterly cross-sectional regressions and this table reports means and medians over the 84 quarters in our sample. We test for significance using the time series variation in the regression coefficients over these 84 periods. We report the significance level based on Newey-West standard errors next to means, and the significance of a Wilcoxon test next to the medians. The asterisks indicate significance at the 1% level (***), 5% level (**), and 10% level (*). V(s)/V(r) is the relative pricing error based on Hasbrouck (1993). |AR30| is the absolute value of the 30-minute quote midpoint return autocorrelation. TOT is the fraction of shares outstanding held by institutions who file quarterly 13F reports. RES is the trade-weighted relative effective spread. QVOL is share trading volume, MV is the market value of equity, and Price is the share price. Lag1 indicates a value lagged by one quarter. Ln is the natural logarithm. Lag1_DV is the lagged value of the respective dependent variable.

Dependent variable	V(s)/V(r)		ln |AR30|		V(s)/V(r)		ln |AR30|	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Intercept	0.804 ***	0.794 ***	-1.989 ***	-1.964 ***	0.508 ***	0.486 ***	-1.827 ***	-1.789 ***
Lag1Tot	-0.064 ***	-0.058 ***	-0.392 ***	-0.394 ***	-0.048 ***	-0.044 ***	-0.360 ***	-0.347 ***
Lag1_DV					0.340 ***	0.398 ***	0.088 ***	0.082 ***
Lag1_RES	0.184	-0.311	11.445 ***	10.849 ***	-0.170	-0.404 **	10.456 ***	10.009 ***
Lag1LnMV	0.013 ***	0.012 ***	0.012	0.003	0.006 ***	0.005 ***	0.007	-0.007
Lag1LnPrice	-0.052 ***	-0.057 ***	-0.163 ***	-0.151 ***	-0.033 ***	-0.036 ***	-0.144 ***	-0.140 ***
Lag1LnQVol	-0.028 ***	-0.024 ***	-0.056 ***	-0.044 ***	-0.013 ***	-0.011 ***	-0.047 ***	-0.036 ***
Number of stocks	1,135	1,110	1,135	1,110	1,135	1,110	1,135	1,110
Adj. R2	0.230	0.229	0.062	0.057	0.382	0.385	0.071	0.068

Table 5: The cross-sectional effect of institutional holdings and changes in holdings on pricing errors

The sample is based on NYSE-listed securities between 1983 and 2003. We conduct quarterly cross-sectional regressions and this table reports means and medians over the 84 quarters in our sample. We test for significance using the time series variation in the regression coefficients over these 84 periods. We report the significance level based on Newey-West standard errors next to means, and the significance of a Wilcoxon test next to the medians. The asterisks indicate significance at the 1% level (***), 5% level (**), and 10% level (*). V(s)/V(r) is the relative pricing error based on Hasbrouck (1993). |AR30| is the absolute value of the 30-minute quote midpoint return autocorrelation. TOT is the fraction of shares outstanding held by institutions who file quarterly 13F reports. TOTChg is the signed percentage of shares outstanding by which these institutions' aggregate holdings have changed during a quarter. RES is the trade-weighted relative effective spread. QVOL is share trading volume, MV is the market value of equity, and Price is the share price. Lag1 indicates a value lagged by one quarter, Lag2 indicates a value lagged by two quarters. Ln is the natural logarithm. Lag1_DV is the lagged value of the respective dependent variable.

Dependent variable	V(s)/V(r)		ln |AR30|		V(s)/V(r)		ln |AR30|	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Intercept	0.802 ***	0.783 ***	-1.992 ***	-2.030 ***	0.504 ***	0.477 ***	-1.830 ***	-1.783 ***
Lag1TOTChg	-0.032 ***	-0.041 ***	-0.369 ***	-0.340 ***	-0.037 ***	-0.047 ***	-0.351 ***	-0.356 ***
Lag2Tot	-0.066 ***	-0.060 ***	-0.407 ***	-0.380 ***	-0.048 ***	-0.040 ***	-0.374 ***	-0.351 ***
Lag1_DV					0.342 ***	0.398 ***	0.088 ***	0.083 ***
Lag1_RES	0.161	-0.355	11.552 ***	10.717 ***	-0.172	-0.384 **	10.543 ***	9.897 ***
Lag1LnMV	0.014 ***	0.014 ***	0.013	0.004	0.006 ***	0.006 ***	0.008	-0.003
Lag1LnPrice	-0.053 ***	-0.057 ***	-0.164 ***	-0.146 ***	-0.034 ***	-0.034 ***	-0.144 ***	-0.129 ***
Lag1LnQVol	-0.028 ***	-0.024 ***	-0.056 ***	-0.044 ***	-0.013 ***	-0.011 ***	-0.047 ***	-0.029 ***
Number of stocks	1,116	1,091	1,116	1,091	1,116	1,091	1,116	1,091
Adj. R2	0.234	0.233	0.064	0.058	0.386	0.392	0.073	0.068

Table 6: The cross-sectional effect of momentum and contrarian changes in institutional holdings

The sample is based on NYSE-listed securities between 1983 and 2003. We conduct quarterly cross-sectional regressions and this table reports means and medians over the 84 quarters in our sample. We test for significance using the time series variation in the regression coefficients over these 84 periods. We report the significance level based on Newey-West standard errors next to means, and the significance of a Wilcoxon test next to the medians. The asterisks indicate significance at the 1% level (***), 5% level (**), and 10% level (*). V(s)/V(r) is the relative pricing error based on Hasbrouck (1993). |AR30| and |AR60| are the absolute values of the 30 and 60-minute quote midpoint return autocorrelation, respectively. TOT is the fraction of shares outstanding held by institutions who file quarterly 13F reports. TOTChg is the signed percentage of shares outstanding by which these institutions' aggregate holdings have changed during a quarter. TOTChgMom includes only increases (decreases) in reported holdings following a positive (negative) buy-and-hold return in the previous quarter, and vice versa for TOTChgCont. RES is the trade-weighted relative effective spread. QVOL is share trading volume, MV is the market value of equity, and Price is the share price. Lag1 indicates a value lagged by one quarter, Lag2 indicates a value lagged by two quarters. Ln is the natural logarithm. Lag1_DV is the lagged value of the respective dependent variable.

Dependent variable	V(s)/V(r)		ln |AR30|		ln |AR60|	
	Mean	Median	Mean	Median	Mean	Median
Intercept	0.504 ***	0.476 ***	-1.820 ***	-1.782 ***	-1.532 ***	-1.700 ***
Lag1TOTChgMom	-0.024 *	-0.045 **	-0.149	-0.141	-0.248 **	-0.142 *
Lag1TOTChgCont	-0.056 ***	-0.048 ***	-0.708 ***	-0.590 ***	-0.480 ***	-0.520 ***
Lag2Tot	-0.048 ***	-0.040 ***	-0.373 ***	-0.352 ***	-0.227 ***	-0.203 ***
Lag1_DV	0.342 ***	0.399 ***	0.088 ***	0.083 ***	0.041 ***	0.038 ***
Lag1_RES	-0.176	-0.392 **	10.456 ***	9.834 ***	11.267 ***	11.726 ***
Lag1LnMV	0.007 ***	0.007 ***	0.008	0.000	-0.006	-0.013
Lag1LnPrice	-0.034 ***	-0.034 ***	-0.147 ***	-0.131 ***	-0.052 **	-0.031 ***
Lag1LnQVol	-0.013 ***	-0.012 ***	-0.047 ***	-0.029 ***	-0.087 ***	-0.075 ***
Number of stocks	1,116	1,091	1,116	1,091	1,116	1,090
Adj. R2	0.387	0.392	0.074	0.068	0.048	0.046

Table 7: The cross-sectional effect of analyst coverage on pricing errors

The sample is based on NYSE-listed securities between 1983 and 2003. We conduct quarterly cross-sectional regressions and this table reports means and medians over the 84 quarters in our sample. We test for significance using the time series variation in the regression coefficients over these 84 periods. We report the significance level based on Newey-West standard errors next to means, and the significance of a Wilcoxon test next to the medians. The asterisks indicate significance at the 1% level (***), 5% level (**), and 10% level (*). V(s)/V(r) is the relative pricing error based on Hasbrouck (1993). |AR30| is the absolute value of the 30-minute quote midpoint return autocorrelation. TOT is the fraction of shares outstanding held by institutions who file quarterly 13F reports. TOTChg is the signed percentage of shares outstanding by which these institutions' aggregate holdings have changed during a quarter. NumAn is the number of analysts covering a stock. RES is the trade-weighted relative effective spread. QVOL is share trading volume, MV is the market value of equity, and Price is the share price. Lag1 indicates a value lagged by one quarter, Lag2 indicates a value lagged by two quarters. Ln is the natural logarithm. Lag1_DV is the lagged value of the respective dependent variable.

Dependent variable	V(s)/V(r)		V(s)/V(r)		ln |AR30|		ln |AR30|		ln |AR60|		ln |AR60|	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Intercept	0.479 ***	0.447 ***	0.477 ***	0.451 ***	-1.932 ***	-1.953 ***	-1.943 ***	-1.970 ***	-1.702 ***	-1.743 ***	-1.703 ***	-1.720 ***
Lag1TOTChg			-0.042 ***	-0.055 ***			-0.349 ***	-0.401 ***			-0.354 ***	-0.287 ***
Lag2Tot			-0.050 ***	-0.042 ***			-0.380 ***	-0.339 ***			-0.231 ***	-0.195 ***
Lag1Tot	-0.050 ***	-0.045 ***			-0.367 ***	-0.344 ***			-0.243 ***	-0.214 ***		
LagLnNumAn*10^2	0.014	0.273	0.055	0.155	-0.466	0.043	-0.588	-0.526	-1.551	-1.749 **	-1.615 *	-2.245 ***
Lag1_DV	0.377 ***	0.423 ***	0.379 ***	0.426 ***	0.085 ***	0.081 ***	0.086 ***	0.081 ***	0.040 ***	0.041 ***	0.040 ***	0.042 ***
Lag1_RES	-0.182	-0.497 **	-0.186	-0.524 **	13.793 ***	13.017 ***	13.899 ***	12.998 ***	14.471 ***	13.876 ***	14.415 ***	13.570 ***
Lag1LnMV	0.005 ***	0.007 ***	0.005 ***	0.006 ***	0.008	0.013	0.009	0.005	-0.004	-0.002	0.000	0.000
Lag1LnPrice	-0.033 ***	-0.037 ***	-0.033 ***	-0.035 ***	-0.126 ***	-0.119 ***	-0.125 ***	-0.114 ***	-0.029	-0.027	-0.034 *	-0.025 *
Lag1LnQVol	-0.011 ***	-0.009 ***	-0.011 ***	-0.010 ***	-0.045 ***	-0.036 ***	-0.045 ***	-0.033 ***	-0.080 ***	-0.064 ***	-0.083 ***	-0.072 ***
Number of stocks	1,089	1,072	1,074	1,056	1,089	1,072	1,074	1,056	1,088	1,072	1,073	1,056
Adj. R2	0.398	0.397	0.402	0.395	0.067	0.063	0.068	0.067	0.044	0.042	0.045	0.044

Table 8: The cross-sectional effect of daily institutional trading on pricing errors

The sample is based on NYSE-listed securities between Jan 1, 2000 and Dec 31, 2003. The sample contains only stocks with at least 100 trades per day. We conduct daily cross-sectional regressions and this table reports means and medians over the 1,003 days in our sample. We test for significance using the time series variation in the regression coefficients over these periods. We report the significance level based on Newey-West standard errors next to means, and the significance of a Wilcoxon test next to the medians. The asterisks indicate significance at the 1% level (***), 5% level (**), and 10% level (*). V(s)/V(r) is the relative pricing error based on Hasbrouck (1993), estimated daily. TOT is the most recent quarterly fraction of shares outstanding held by institutions who file quarterly 13F reports, as reported in the most recent quarterly report. Ivol/TotVol is the fraction of total NYSE volume due to institutional trading. Ivol/Shrout is NYSE institutional volume divided by the number of shares outstanding. Contrarian trades are buys (sells) when the return on the previous day is negative (positive), and vice versa for momentum trades. RES is the trade-weighted relative effective spread. MV is the quarterly market value of equity, and VWAP is the volume-weighted average share price. Lag1 indicates a value lagged by one quarter. Ln is the natural logarithm. Lag1_DV is the lagged value of the respective dependent variable. All variables except TOT and MV are measured daily.

Panel A: Aggregate institutional trading volume

Dependent variable V(s)/V(r)	Mean	Median	Mean	Median
Intercept	0.8131 ***	0.7908 ***	0.8012 ***	0.7813 ***
LagIvol/TotVol			-0.0331 ***	-0.0317 ***
LagIvol/Shrout	-0.0019 ***	-0.0017 ***		
LagTot	-0.0322 ***	-0.0310 ***	-0.0424 ***	-0.0412 ***
Lag1_DV	0.1918 ***	0.1905 ***	0.2003 ***	0.2004 ***
LagDlyRES	10.4554 ***	10.4122 ***	7.8133 ***	8.3578 ***
LagLnMV	-0.0269 ***	-0.0260 ***	-0.0242 ***	-0.0232 ***
LagLnVWAP	-0.0032	-0.0024 ***	-0.0078 ***	-0.0062 ***
Number of stocks	351	351	351	351
Adj. R2	0.176	0.171	0.170	0.164

Panel B: Institutional trading volume conditional on trade direction and previous-day returns

	Mean	Median	Mean	Median
Intercept	0.8143 ***	0.7876 ***	0.7992 ***	0.7833 ***
LagIvol/TotVol (contrarian buy)			-0.0304 ***	-0.0289 ***
LagIvol/TotVol (momentum buy)			-0.0315 ***	-0.0333 ***
LagIvol/TotVol (contrarian sell)			-0.0177 ***	-0.0212 **
LagIvol/TotVol (momentum sell)			-0.0226 ***	-0.0268 ***
LagIvol/Shrout (contrarian buy)	-0.0032 ***	-0.0032 ***		
LagIvol/Shrout (momentum buy)	-0.0026 ***	-0.0021 ***		
LagIvol/Shrout (contrarian sell)	-0.0014 **	-0.0016 ***		
LagIvol/Shrout (momentum sell)	-0.0007	-0.0003		
LagTot	-0.0315 ***	-0.0304 ***	-0.0425 ***	-0.0413 ***
Lag1_DV	0.1914 ***	0.1894 ***	0.2002 ***	0.1985 ***
LagDlyRES	10.4798 ***	10.0130 ***	7.7796 ***	7.6278 ***
LagLnMV	-0.0270 ***	-0.0261 ***	-0.0243 ***	-0.0231 ***
LagLnVWAP	-0.0031	-0.0019 ***	-0.0075 ***	-0.0060 ***
Number of stocks	351	351	351	351
Adj. R2	0.185	0.181	0.178	0.174

Figure 1: Time-series and cross-sectional distribution of relative pricing errors

The sample is based on NYSE-listed securities between 1983 and 2003. This figure shows the mean, median, 25th percentile, and the 75th percentile of the cross-sectional distribution in each quarter. The vertical bars indicate one cross-sectional standard deviation in each direction. V(s)/V(r) is the relative pricing error based on Hasbrouck (1993).

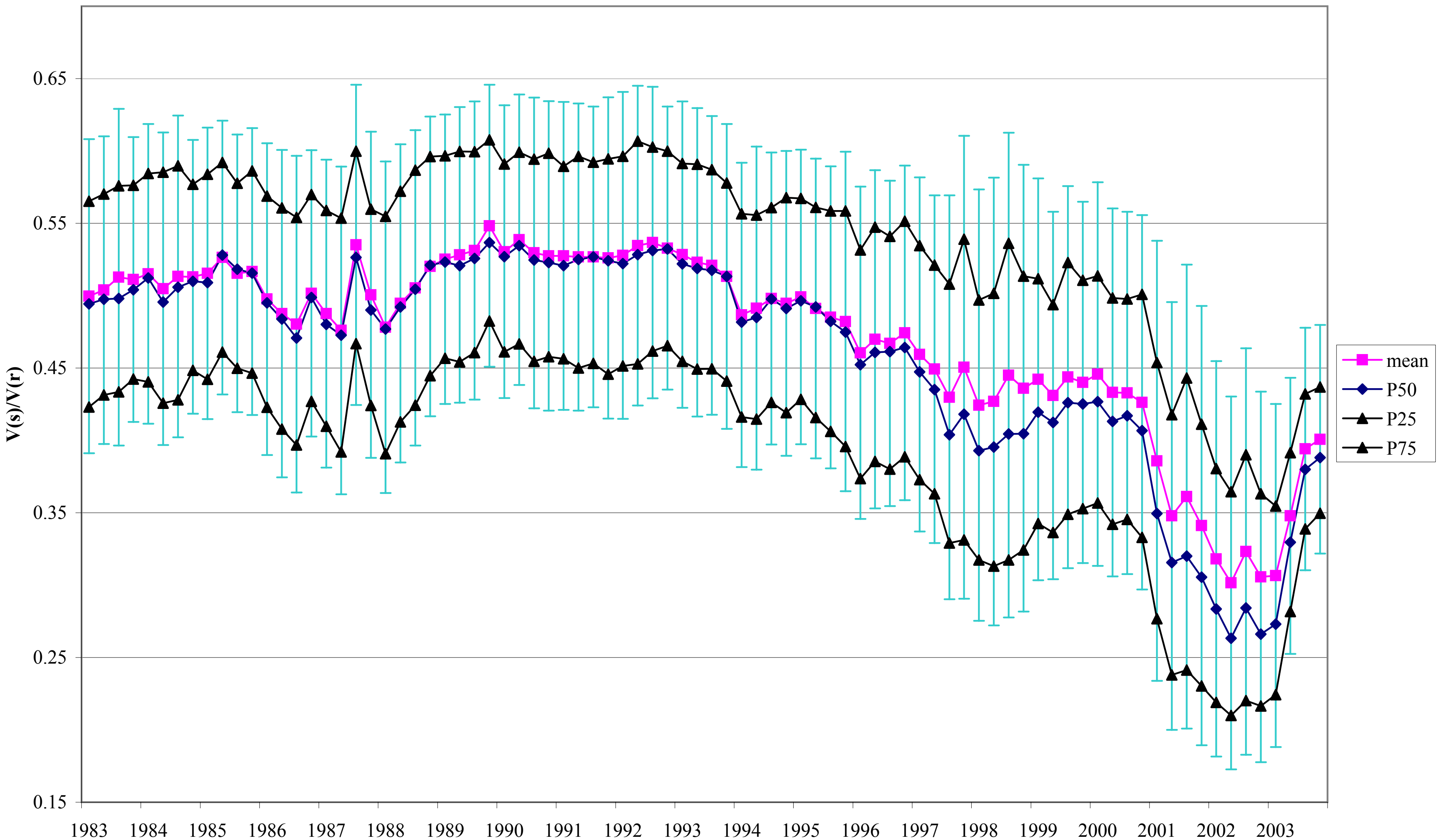
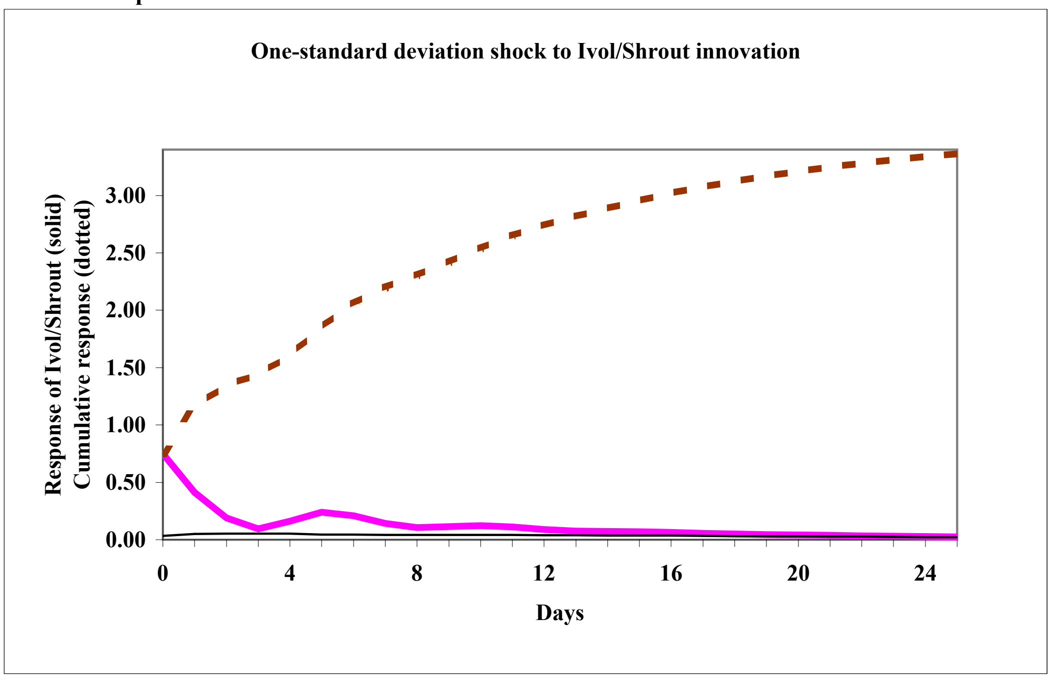
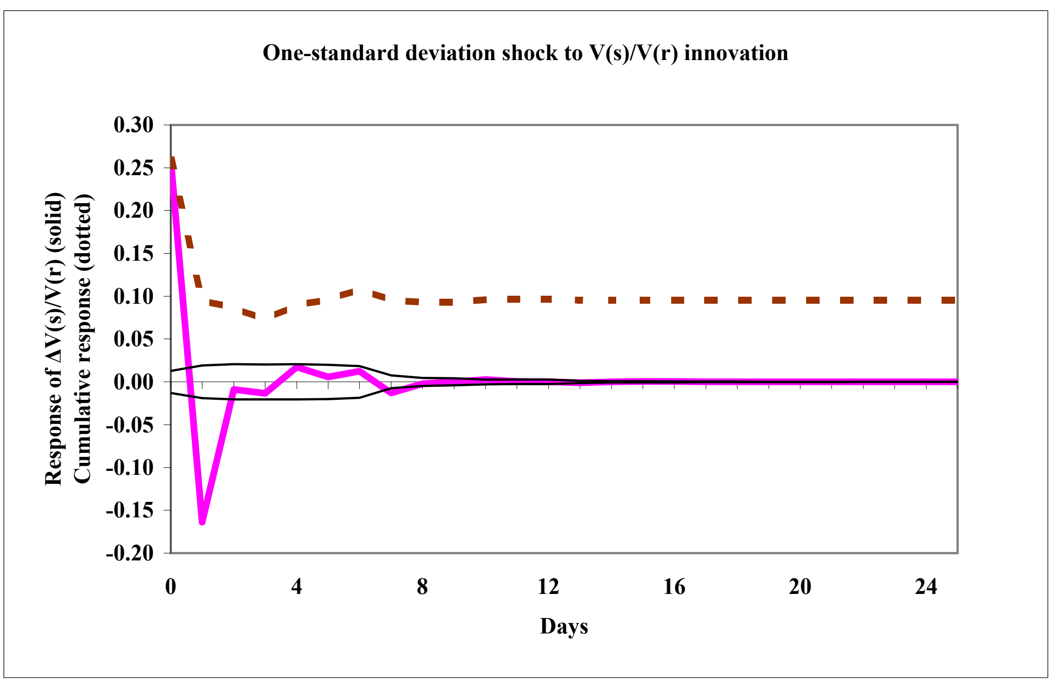


Figure 2: Dynamics of daily institutional turnover and pricing errors

The sample contains, on average across days, 351 NYSE-listed stocks between Jan 1, 2000 and Dec 31, 2003. The figure shows orthogonalized impulse response function of a two-equation five-lag structural VAR model that relates the first difference of daily pricing errors (ΔV(s)/V(r)) to daily institutional turnover, Ivol/Shrout. Both series represent the cross-sectional average on each trading day, and are standardized using their time-series mean and standard deviation. The contemporaneous effect of V(s)/V(r) on Ivol/Shrout is assumed to be zero in the structural model. The solid line represents the impulse response and the dotted line represents the cumulative response. The thin lines plot two standard errors of the impulse response around zero.

Panel A: Responses to shocks in own innovations





Panel B: Responses to shocks in cross-innovations