Risk Sharing, Costly Participation, and Intermediation

Terrence Hendershott
U.C. Berkeley

Sunny X. Li
VU University Amsterdam

Albert J. Menkveld
VU University Amsterdam

Mark S. Seasholes
HKUST

This Version: 15-May-2010*

Abstract

We study two groups of traders previously shown to provide immediacy on the NYSE. Consistent with being compensated for providing liquidity, market makers' inventories (and individuals' net trades) are negatively correlated with past/current returns and positively correlated with future returns. We estimate a state-space model at a monthly frequency and show the groups' positions and net trades are associated with 1.63% and 1.66% of monthly volatility, respectively. The results are larger for smaller stocks (2.62% and 1.95%). When the two groups trade in concert, their positions/trades explain 27.15% of transitory variance. Our results are consistent with a model of imperfect risk sharing and costly participation. Market makers have low costs and continuously participate in the market. After building positions, they share/spread risk and unwind positions by trading with individuals (whose participation costs are higher).

Keywords: Liquidity, Transitory Volatility

JEL Number: G12, G14

*Contact information: **Terrence Hendershott**, Haas School of Business, University of California at Berkeley, 545 Student Services Bldg 1900, Berkeley CA 94720, Tel: +1 510.643.0619, Email: hender@haas.berkeley.edu; **Sunny X. Li**, VU University Amsterdam, FEWEB, De Boelelaan 1105, 1081 HV, Amsterdam, Netherlands, Email: xysunnyli@gmail.com; **Albert J. Menkveld**, VU University Amsterdam, FEWEB, De Boelelaan 1105, 1081 HV, Amsterdam, Netherlands, Tel: +31.20.598.6130, Email: albertjmenkveld@gmail.com; **Mark S. Seasholes**, HKUST Dept of Finance (Rm 2413), Clear Water Bay, Kowloon Hong Kong, HK Tel: +852.2358.7668, USA Tel: +1.510.931.7531, Email: Mark.Seasholes@emailias.edu

1 Introduction

Many empirical asset pricing studies use monthly stock prices and returns. How "noisy" are these data? More precisely, how large are transitory price deviations around fundamental values? In other words, do liquidity demanding trades *push* prices away from fundamentals? How can we use trading measures from market participants to help answer these questions?

This paper provides answers to the questions posed above. We start with the hypothesis that a stock's observable market price reflects both information about stock's fundamental value (also known as its "efficient price") and the effects of transitory liquidity shocks (also known as "price pressure"). Neither the efficient, nor the transitory, price is observable. Therefore, our first step is to estimate a state space (statistical) model and employ a Kalman filter to disentangle changes in a stock's fundamental value from transitory price pressure.

Our first result is that, at a monthly frequency, transitory price pressure is more than one quarter the magnitude of efficient price variation. The amount is both statistically and economically significant. The Kalman filter provides time series estimates of a stock's efficient prices and associated price deviations. We are then able to study correlations of changes in efficient prices, price deviations, and trading measures from different market participants.

Why study the trading behavior of different market participants? Our hypothesis that observable prices are temporarily pushed away from fundamental values is consistent with (temporarily) incomplete risk sharing. Suppose a stock is trading at its fundamental value. When a group of traders demands immediacy (wants to trade), there may not be enough market makers (liquidity suppliers) to absorb the sell orders or fill the buy orders. Prices adjust as in Grossman and Miller (1988) in order to compensate market makers for supplying liquidity.

One result of limited risk-bearing capacity is that financial econometricians should observe correlations between market makers' trading, contemporaneous price movements, and price mean reversion. Market makers buy (sell) as prices fall (rise). The initial downward (upward) price pressure is then followed by prices rising (falling). The mean reversion in prices provides the compensation necessary to induce some market participants to supply liquidity.

In this paper, we use New York Stock Exchange (NYSE) specialists' inventories as a (negative) proxy for liquidity demanding trades. That is, market makers tend to buy (sell) when impatient investors desire to sell (buy). We also study the net trades of a second group that Kaniel, Saar, and Titman (2008) show to trade against prices—mainly, individuals on the NYSE. Our trading variables are correlated with Kalman filter estimates of stocks' efficient price changes and price deviations. We therefore add the trading variables to the state state model in order to better understand transitory price pressure.

Including trading variables into the state space model produces a plethora of results. We briefly summarize the findings here (while noting that results 4, 5, and 6 below represent this paper's main contributions):

1. When focusing on idiosyncratic returns, specialists' inventories (and individuals' net trades) are negatively correlated with past/current returns. Both groups can be said to trade against price movements.

2. Specialists and individuals are (partially) compensated for providing immediacy via return reversals. Their respective inventories and net trades are positively correlated with future returns.

3. Specialists' inventories are negatively auto-correlated indicating this group manages inventory risk in a manner consistent with traditional models of market making. Individuals' net trades are positively auto-correlated. There is no evidence that individuals mean revert their positions nor do they appear to manage inventory risk. In brief, individuals do not act in a manner consistent with traditional models of market making.

4. Specialists' inventories are positively correlated with individuals' future net trades. This result is consistent with specialists sharing/spreading risk and unwinding their positions by trading with individuals.

5. A one standard deviation in specialists' inventories is associated with a 1.63% deviation in transitory prices. A one standard deviation in individuals' net trades is associated with 1.66% deviation in transitory prices. The results are larger for small stocks (2.62% and 1.95% respectively.) Specialists' inventories and individuals' net trades explain 12% to 13% of transitory price variance, respectively.

6. Our two trading variables provide separate and complementary information about liquidity demands. When both variables are included in our state-space model, the trading variables explain 27.15% of transitory variance. Put differently, downward (upward) transitory price pressure is particularly severe when specialists are long (short) and individuals are buying (selling).

Our results are consistent with a model of imperfect risk sharing and costly participation. Market makers have low participation costs and continuously monitor the market. They are able to quickly trade against price movements. Individuals, on the other hand, have higher participation costs and participate intermittently. Some individuals trade against initial price movements; others delay their trades. When market makers unwind positions, they are able to trade with the second (delayed) group of individuals.

1.1 Related Literature

Our paper is linked to three strands of recent literature. First, studies of NYSE specialists date back to Madhavan and Smidt (1993) and have recently been enhanced by Hendershott and Seasholes (2007) and Hendershott and Menkveld (2010). As with these studies, we show specialists trade against price movements and they work to mean revert their inventories.

Second, a paper by Kaniel, Saar, and Titman (2008) is our motivation for studying individuals' net trades on the NYSE. The authors study a large cross-section of NYSE stocks from January 2000 to December 2003. Individuals are shown to buy stocks that have recently fallen in price and they sell stocks that have recently risen. Sorting stocks by the degree of buying and selling allows the authors to form a long-short portfolio that earns over 120 basis points in the 20 days following formation.

Third, two recent introduce an econometric approach to disentangling permanent and transitory price effects. We employ such an approach in this paper. Menkveld, Koopman, and Lucas (2007) and Hendershott and Menkveld (2010) use state space models. The second paper studies prices at a daily frequency.

Our paper is structured as follows. Section 2 outlines an economic framework with imperfect risk sharing and costly participation. Section 3 describes the paper's data and provides overview statistics. Section 4 estimates a base case version of the state space model. This version does not use trading variables. Section 5 introduces trading variables into the state space model. Section 6 concludes.

2 Theoretical Framework

The Economy: We assume a simple world with three dates denoted $t = \{1, 2, 3\}$ and two assets. The first asset is a riskless security, used as the numeraire good, and assumed to have a zero rate of return. The second is a risky asset that pays $\tilde{D}$ units of the consumption good at t=3, where $\tilde{D} = \bar{D} + \tilde{\epsilon}_1 + \tilde{\epsilon}_2$. The distribution of $\tilde{\epsilon}_t$ is normal with mean 0 and variance σ_t^2. We denote $\tilde{P}_t$ as the risky asset's price on date t, with $\tilde{P}_3 = \tilde{D}$.

Participants: There are three types of participants in the market and all assumed to be present with measure zero. For concreteness, consider Group a to be a long-term investor called "institutions", Group b to be a long-term investor called "individuals", and Group m to be a short-term investor called "market makers" or "arbitrageurs".

All groups have an initial endowment θ of the risky asset. Groups a and b have opposite exposure to a non-traded risk which is perfectly correlated with the t=3 payoff of the risky asset.[1] Group a investors receive an endowment $+(\tilde{D} - \bar{D})$ of the consumption good and Group b receives an endowment of $-(\tilde{D} - \bar{D})$.

Participation Costs: Group a's participation costs are zero, so they can trade freely at both t=1 and t=2. Group b has a participation cost of κ at t=1. Due to this cost, some individuals refrain from trading at t=1 leading to a "participation intensity" that is denoted λ. At t=2, all individuals participate. Importantly, λ is endogenously determined in this model.[2]

Timing: At t=2, all three groups have acquired some information about $\tilde{D}$ from the realization ϵ_1. We define the expectation of $\tilde{D}$ at t=2 as $\mathbb{E}_2\left[\tilde{D}\right] \equiv \bar{D} + \epsilon_1$. At t=1, orders in the market may not be balanced due to the delay of $(1-\lambda)$ investors from Group b. Market makers offset temporary imbalances by taking positions that they hold until t=2. At t=2, all individual investors are present in the market and market makers are able to unwind their positions.

Agents' Maximization Problems Investors work to maximize their expected utilities of wealth at t=3 which is denoted $\mathbb{E}\big[U(W_3^j)\big]$ for group j. We assume agents have exponential utility functions of the form $U(W_3^a) = -e^{\alpha W}$. Let $\bar{x}_t^j$ be the number of risky asset units owned by group j at date t. The group's excess demand is denoted $x_t^j = \bar{x}_t^j - \theta$. For example, at t=1, institutions own $x_t^a + \theta$ of the risky asset. We also use B_t^j to denote the respective holdings of the riskfree asset. Wealth at time t is given by $B_t^j + \tilde{P}_t x_{t-1}^j$. We solve for equilibrium holdings and price by backwards induction. Please see Appendix A for associated proofs and expanded equations.

Equilibrium Prices and Net Trades: We present equilibrium prices and holdings in the chart below. The term "Group b(p)" indicates individuals who choose to participate at both t=1 and t=2. The term "Group b(np)" indicates individuals who only participate at t=2.

[1]Since these endowments are perfectly correlated with risky asset payoff, the non-traded risk encourages participants to trade in order to share risk. For examples of papers with a related mechanism, see also Lo, Mamaysky, and Wang (2004) and Vayanos and Wang (2009).

[2]We can consider that individual investors may not participate at short term horizons due to participation costs but in the long run they will all participate at least once in the market. Our 3-date model is stylized. Dates t=1 to t=2 can be thought of as a short-term horizon while dates t=2 to t=3 represent a longer-time horizon.

	t=1	t=2
Price of Risky Asset	$\bar{D} - \theta\alpha(\sigma_1^2 + \sigma_2^2) - \frac{1-\lambda}{2+\lambda}$	$\bar{D} + \epsilon_1 - \theta\alpha\sigma_2^2$
Holdings of Group a	$-\frac{2\lambda+1}{\lambda+2} + \theta$	$-1 + \theta$
Holdings of Group b(p)	$\frac{3}{\lambda+2} + \theta$	$+1 + \theta$
Holdings of Group b(np)	$+\theta$	$+1 + \theta$
Holdings of Group m	$\frac{1-\lambda}{\lambda+2} + \theta$	$+\theta$

Notice that prices equal expected future payoffs minus a risk premium. At t=1, the risk premium is sum of three terms: $\theta\alpha\sigma_1^2 + \theta\alpha\sigma_2^2 + \frac{1-\lambda}{2+\lambda}$. Note that all three terms are ultimately subtracted from expected payouts to get to P_1. The first two terms come from dividend uncertainty. At t=2, the risk premium is $\theta\alpha\sigma_2^2$ reflecting the remaining dividend uncertainty.

Individual investors are indifferent between delaying one period and participating at both t=1 and t=2 when $\mathbb{E}\big[U(W^b(p))\big]=\mathbb{E}\big[U(W^b(np))\big]$. Using this condition, we solve for the endogenously determined participation intensity:

$$\lambda = \frac{3\sqrt{\alpha}}{\sqrt{2\kappa}}\sigma_1 - 2$$

2.1 Model's Results and Predictions

1. We define price pressure as $s_1 \equiv P_1 - \bar{D}$. We find price pressure is negatively related with market makers' inventory positions and negatively related with individual investors' trading positions: $P_1 - \bar{D} = -\alpha\sigma_2^2\theta - \alpha\sigma_1^2\bar{x}_1^m = -\alpha(\sigma_1^2 + \sigma_2^2)\theta - \alpha\sigma_1^2\frac{1-\lambda}{3}x_1^b(p)$.
2. Individual investors' (average) net trading imbalances are persistent. The average net trading imbalance at time t=1 is $\Delta x_1^b = \frac{3\lambda}{2+\lambda}$ and the average net trading imbalance at time t=2 is $\Delta x_2^b = \frac{2(1-\lambda)}{2+\lambda}$.
3. Market makers' inventories are mean-reverting and they unwind t=1 positions with the individual investors who trade at t=2.
4. Market makers' inventories at t=1 are positively correlated with individual investors' net trading imbalances at both t=1 and t=2.
5. Individual investors' participation intensity increases with the risk-aversion parameter (α), uncertainty about future dividends (σ_t), initial endowment of risky assets (θ), and amount of their exposure to the non-tradable risk factor (set to -1 in this paper).

6. If participation costs κ are greater than $\frac{9}{8}\alpha\sigma_1^2$, no individual investor participates at t=1. If κ is less than $\frac{1}{2}\alpha\sigma_1^2$, all individual investors participate at t=1 and t=2.

3 Data and Overview Statistics

We study monthly trading activity and stock prices starting January 1999 and ending December 2005 for a total of 84 months. Four sources provide the data used in this paper.

1. An internal New York Stock Exchange ("NYSE") database called the Specialist Summary File (or "SPETS") contains specialists' closing inventory positions for each stock at the end each month. The NYSE assigns one specialist per stock and a given specialist is responsible for approximately ten stocks.
2. An internal NYSE database called the Consolidated Equity Audit Trail Data (or "CAUD") contains the number of shares bought and sold by individual investors, for each stock, over each month. In addition, the CAUD database provides trading volume. See Kaniel, Saar, and Titman (2008) for further discussion of the CAUD database.
3. The Trades and Quotes ("TAQ") database provides closing midquotes prices. Prices and returns in this paper are measured at the midquote to avoid bid-ask bounce. All prices are adjusted to account for stock splits and dividends.
4. The Center for Research in Security Prices ("CRSP") provides the number of shares outstanding (used to calculate market capitalizations) and information necessary to adjust prices for stock splits/distributions.

We start the 2,357 common stocks common stocks that can be matched across the NYSE, TAQ, and CRSP databases. We construct a balanced panel of data to ensure results are comparable throughout time. There are 1,037 stocks that exist for all 84 months in our sample period. Stocks with an average share price of less than US$ 5 or larger than US$ 1,000 are removed from the sample. The final sample consists of 1,019 stocks.

We convert specialists' inventory positions and individual net trades to US dollars (both variables are originally in numbers of shares.) For a given stock, we multiply the number of shares by the stock's sample average price so as not to introduce price changes directly into the trading variables. Using end-of-month prices would infect/negate the trading variables role as explanatory variables for transitory price effects in the econometric models in Section 4 and 5.

Finally, many results are presented after sorting stocks into size quintiles. To ensure the quintiles have constant compositions throughout the sample period, stocks are ranked based on their average market capitalizations over the entire sample period.

3.1 Summary Statistics

Table 1, Panel A presents summary statistics for seven "raw" variables. For each of the five market capitalization quintiles, we calculate each variable's average value. The smallest quintile's average market capitalization is US$ 0.26 billion while the largest quintile's is US$ 33.90 billion. The last column shows the within standard deviation is US$ 6.57 billion.

[Insert Table 1]

The table also shows overview statistics for trading volume (in millions of shares) and closing mid-quote prices. Trading variables include specialists' inventories (in both thousands of shares and dollars) and individuals' net trades. On average, specialists hold half a million U.S. dollars of inventory for large capitalization stocks. The positive average inventory values may be due to asymmetric costs. Shorting may be more expensive than holding stocks long. The within standard deviation is US$ 1.32 million and substantial relative to the average position. The large standard deviation suggests that specialists are active intermediaries.

Individuals' average net trades are negative across all size quintiles indicating that individuals' positions have been reduced over our sample period. Individual investors, on average, sell US$ 0.20 million in small-cap stocks and US$ 14.12 million in large-cap stocks each month. The within standard deviation of individual net trades is US$ 18.17 million, which is also large relative to their net trades. The summary statistics suggest individual investors participate actively at a monthly frequency.

3.2 Idiosyncratic Variables

Risks associated with market-wide return shocks can be hedged using highly-liquid index products. In addition, aggregate inventories of liquidity providers are exposed to market-wide return shocks. Therefore, our empirical analysis focuses almost entirely on idiosyncratic components of our variables. For each return and trading variable, we construct a common factor equal to the monthly market capitalization weighted average of the underlying variable. We regress each variable on its common factor and save the residual as the corresponding idiosyncratic variable. This procedure is detailed in Appendix F. For notational simplicity, we omit any subsricpts or superscripts referring to "idiosyncratic," and use $Spec_{i,t}$ to denote the idiosyncratic portion of the specialist's dollar inventory (for example).

Table 1, Panel B provides summary statistics for idiosyncratic trading variables used in this paper. Since the idiosyncratic variables are defined as residuals from a market model regression, means are zero. See Appendix F. Therefore, the panel reports standard deviations for the five size quintiles and for the sample as a whole. We see the largest stocks have volatile inventories (2,623.4 in thousands of dollars) and volatile net trades by individuals (39,505.0 also in thousands of dollars).

For completeness, we also report the standard deviations of idiosyncratic returns. As shown in Appendix F, $r_{i,t}^{idio}$ is stock i's residual from a regression on current and lagged innovations in the common return factor. Small stocks have an standard deviation of 13.74% while large stocks have a standard deviation of 10.31%.

3.3 Unit Root Tests

We test for mean-reversion of the specialists' and individuals' inventory positions. The augmented Dickey-Fuller test is performed on a stock-by-stock basis using the regression in the equation below. While the regression to test specialists' inventories is shown below, a similar expression is used to test individuals' net trades. Note, $\Delta Spec_{i,t}$ is the first difference of $Spec_{i,t}$. The latter is discussed above and both are defined explicitly in Appendix F.

$$\Delta Spec_{i,t} = \alpha + \beta Spec_{i,t-1} + \phi_1 \Delta Spec_{i,t-1} + \ldots + \phi_4 \Delta Spec_{i,t-4} + \varepsilon_{i,t}$$

Table 2 presents the results of the augmented Dickey-Fuller tests. The table reports the cross-sectional mean of the t-statistic associated with the β coefficient. The table also reports the p-value of a meta test statistic that counts the number of significant t-values under 10% critical value of augmented Dickey-Fuller test.[3] This meta test statistic is binomially distributed under null where the probability of "success" equals the significance level of the augmented Dickey-Fuller test performed for each stock estimation. We use a 10% critical value if the cross-sectional mean is negative and a 90% value if the cross-sectional mean is positive.

[Insert Table 2]

[3]The 10% critical value is -2.57, see Cheung and Lai (1995).

We reject the existence of unit roots in the specialists' inventory positions at all conventional levels. A total of 862 of the 1,019 stocks reject the null. Our results indicate that NYSE specialists behave in a manner consistent with theoretical models of market making. After building a position, specialists quickly undo their trades and mean-revert inventories towards target levels.[4]

We fail to reject the existence of unit roots in the individual inventory positions. Cross-sectionally, we fail to reject for 968 of the 1,019 stocks. Our results indicate that, at the aggregate level, individuals do not mean revert their holdings. The numbers of significant $t-$values are available in the posted supplementary material.[5]

NYSE specialists' inventory levels are stationary, while the levels for individuals are not stationary. Therefore, and throughout the paper, we use the *level* of NYSE specialists' inventories ($Spec_{i,t}$) and the *change in levels*, or net trades, of individuals' holdings ($\Delta Indv_{i,t}$).

4 Price Decomposition—Base Case

This section presents the base case version of our state space (statistical) model. The base case only uses price and innovations to prices (i.e., no trading variables) when decomposing a stock i's (observed) price into two unobserved components. The first is called the "efficient price" and reflects the stock's fundamental value. In our statistical model, the efficient price follows a martingale process and is denoted $m_{i,t}$ throughout the paper. The second component is called the "transitory price" and represents price pressure. The transitory component is stationary and is denoted $s_{i,t}$.

The model is estimated on a stock-by-stock basis. For estimation purposes, we use stock i's log price, expressed in basis points, after removing a required return. We denote this price as $p_{i,t}$. The required return is equal to the monthly risk free rate plus the stock's beta times a market risk premium of 6%. Appendix F provides details on the associated calculations. All values are at the end of a given month t. The base case state space model consists of the

[4] For examples, see Ho and Stoll (1981), Madhavan and Smidt (1993), and Grossman and Miller (1988).

[5] See http://dl.dropbox.com/u/5179651/supplementarytables.pdf. We also test individual inventories using 5% critical values. The 10% threshold represents a weaker-than-normal test, that still ends with 968 out of 1,019 stocks failing to reject.

following three equations:

$$p_{i,t} = m_{i,t} + s_{i,t} \quad (1)$$

$$m_{i,t} = m_{i,t-1} + \beta_i f_t + w_{i,t} \quad (2)$$

$$s_{i,t} = \phi_i s_{i,t-1} + \beta_{i,0} f_t + \ldots + \beta_{i,3} f_{t-3} + \epsilon_{i,t} \quad (3)$$

Above, f_t represents the innovation of a common (market-wide) factor. Additional information about calculating f_t is given in Appendix F. The idiosyncratic innovation of stock i's efficient price is denoted $w_{i,t}$ and is one focus of this paper since it represents undiversifiable risk. The f_t terms in the transitory price equation captures current and lagged adjustment to common factor innovation. $\epsilon_{i,t}$ is the idiosyncratic innovation of the stock's transitory price pressure.

Both $w_{i,t}$ and $\epsilon_{i,t}$ are assumed to be normal and independent in the base case estimation. We revisit this assumption in Section 5 when we include trading variables in the state space model. Price pressure persistence is accounted for by the auto-regressive term in the transitory ($s_{i,t}$) equation. The AR(1) coefficient, ϕ_i, is constrained to be nonnegative.

The model is estimated with maximum likelihood and exploits a Kalman filter. The estimation is implemented in Ox using standard optimization routines. The Kalman filter routines are from `ssfpack` which is an add-on package in Ox. See Koopman, Shephard, and Doornik (1999) for additional information about related estimation procedures. The optimization procedure follows steps designed to avoid getting stuck in local maxima. Appendix B has additional details.

There are at least three advantages associated with using a state space model. First, maximum likelihood estimation is asymptotically unbiased and efficient. Second, the statistical model offers a structural analysis that helps identify effects that would otherwise be unobserved. After estimation, the Kalman filter offers an in-sample decomposition of price time series into the efficient and transitory components. The decomposition is available at any point in the sample period using past and current prices. Third, the Kalman filter helps to deal with missing observations in a simple way without losing information. The model implies the differenced price series ($\Delta p_{i,t}$) follows a MA(1) process which can be expressed as an infinite lag autoregressive model or AR(∞). It is cumbersome to estimate such a model if the price series has missing values. The Kalman filter in the state space model considers the likelihood of all level series changes even if they have missing observations over multiple periods. Methods based on differenced series do not consider such information.

Table 3 presents the base case estimates. Looking at all stocks, we see an estimated value of 849 basis points for the standard deviation of $w_{i,t}$. Transitory shocks are persistent as shown by the 0.35 value of ϕ_i. The fourth column shows the total standard deviation associated with transitory shocks. We calculate the total as $\left(\frac{\sigma(\epsilon)^2}{1-\phi_i^2}\right)^{\frac{1}{2}}$ and find it is equal to 452 basis points on average. The value of 452 basis points is key to analysis in Section 5 where we use the value as a point of comparison for the amount of idiosyncratic volatility that can be explained by our trading variables.

[Insert Table 3]

Table 3 shows the expected result that fundamental volatility is higher in smaller stocks. For the smallest quintile, $\sigma(w)$ is 1,059 bp; for the largest quintile, $\sigma(w)$ is 691 bp. Interestingly, the autocorrelation coefficient, ϕ_i, does not vary significantly across size quintiles.

To assess the economic importance of the (total) transitory shock, we calculate the ratio of transitory variance to efficient price variance. Using the numbers on the top row of Table 3, we see $\frac{452^2}{849^2} = 28.34\%$ suggesting price pressure is economically large at a monthly frequency. The finding that transitory price variance is 28.34% the magnitude of efficient price variance at a monthly frequency represents the first contribution of this paper.

4.1 Correlations of Trading and Price Variables

We use the estimated coefficients from Table 3 and a Kalman filter to extract estimated changes to a stock's efficient price as well as the estimated price pressure. We can also calculate the idiosyncratic portion of a stock's returns. We then correlate these return/price variables with the idiosyncratic portion of specialists' inventories and individuals' net trades.

The Kalman filter gives the conditional expectation of the efficient price of stock i at time t. We denote this quantity as $\hat{m}_{i,t} = E\left[m_{i,t}|\mathcal{P}_{i,t}\right]$ where $\mathcal{P}_{i,t}$ represents the set containing all current and past prices.[6] We define the return (or change) of the efficient price and the price pressure of stock i for month t as:

$$\begin{aligned} \Delta\hat{m}_{i,t} &= E\left[m_{i,t}|\mathcal{P}_{i,t}\right] - E\left[m_{i,t-1}|\mathcal{P}_{i,t-1}\right] \\ \hat{s}_{i,t} &= p_{i,t} - \hat{m}_{i,t} \end{aligned}$$

We calculate a correlation matrix for each stock, across five variables, using time t and lagged values, and using all 84 months of data. The five variables are: $Spec_{i,t}$, $\Delta Indv_{i,t}$,

[6] Additional details about the Kalman filter are provided in Appendix C.

$r_{i,t}^{idio}$, $\Delta\hat{m}_{i,t}$, and $\hat{s}_{i,t}$. Appendix F provides details for all five variables. For stock i, the lag-j correlation of two variables $\left\{x_{i,t}^{a}, x_{i,t}^{b}\right\}$ is defined below. At lag j=0, the expression below shows contemporaneous correlations.

$$\rho^{a,b}(j) = \frac{Cov(x_{i,t}^{a}, x_{i,t-j}^{b})}{\sigma_{x_{i,t}^{a}}\sigma_{x_{i,t}^{b}}}$$

Table 4 reports average correlation results—averaged across all stocks' correlation matrices. There are six main results we focus on. First, we note the negative autocorrelation of idiosyncratic returns confirms our conjecture about transitory price deviations. The first-order auto-correlation coefficient is -0.06 and can be found by looking under "All" stocks 3^{rd} row (r^{idio}) and under "Lag 1" 3^{rd} column (r_{-1}^{idio}). The second-order auto-correlation coefficient of r^{idio} is -0.03 as shown under "Lag 2".

[Insert Table 4]

Second, using the first and second order autocorrelations of r^{idio} (-0.06 and -0.03), we can roughly calculate the implied ratio of transitory volatility and permanent volatility. The estimate is 31.58% and compares to the estimated value of 28.34% in the base-case state space model (see the end of the previous section.) Appendix E provides details of our implied ratio calculation. The estimate of 31.58% is important because its shows that, although autocorrelation coefficients appear small in magnitude (-0.06 and -0.03), they can still explain a large fraction of permanent volatility (31.58%).

Third, individuals' net trades are persistent as shown by the 0.31 first order autocorrelation coefficient. As discussed earlier, the positive autocorrelation indicates individual investors holdings do not mean-revert. To further support the finding, note that the second-order auto-correlation coefficient is 0.16 for $\Delta Indv_{i,t}$.

Fourth, individuals' net trades are positively correlated with subsequent idiosyncratic returns (0.03) and negative correlated with previous (-0.17) and current idiosyncratic returns (-0.17). We can infer that the individual investors buy stocks when prices are falling and sell later when prices go up. The price rise appears smaller in magnitude (0.03) than the current fall in prices (-0.17).

Fifth, the correlation between specialist inventories and subsequent idiosyncratic returns is positive (0.05). The negative contemporaneous correlation between price pressure and specialist inventories (-0.09), suggests that the deviations from the specialists' optimal inventory positions are partially compensated by the temporary price deviation.[7] There is a negative contemporaneous correlation between efficient price change and specialist inventory (-0.23).

Sixth, the specialists' inventory positions are positively correlated with subsequent individual net trades (0.07). This finding suggests that specialists unwind their positions by selling to individual investors.

We finish this section by emphasizing the general take-away from Table 4: The correlations between trading variables and price variables provide support for adding trading variables into the state space model.

5 Trading Variables and Price Decomposition

This section uses measures of specialists' and individuals' trading as explanatory variables in the state space model. We begin by including only one group's trading variables at a time. Section 5.1 uses specialists' trading variables and Section 5.2 uses individuals' trading variables. Section 5.3 ends our analysis by presenting a statistical model that simultaneously considers both groups' trading variables. The full state space model with both group's trading variables is given by:

$$
\begin{aligned}
p_{i,t} &= m_{i,t} + s_{i,t} \\
m_{i,t} &= m_{i,t-1} + \beta_i f_t + w_{i,t} \\
w_{i,t} &= \kappa_i^{spec} \tilde{S}pec_{i,t} + \kappa_i^{indv} \Delta \tilde{I}ndv_{i,t} + u_{i,t} \qquad (4) \\
s_{i,t} &= \alpha_i^{spec} Spec_{i,t} + \alpha_i^{indv} \Delta Indv_{i,t} + \alpha_i^D D_{i,t} + \beta_{i,0} f_t + \ldots + \beta_{i,3} f_{t-3} + \epsilon_{i,t} \qquad (5)
\end{aligned}
$$

Because Table 4 shows trading variables are negatively correlated with changes to the efficient price, we add the trading variables to the innovation of the efficient price in Equation (4). Our goal is to avoid a potential omitted variable bias. The tilde over the first letter of a variable indicates autocorrelation has been removed using an AR(1) regression. Appendix F provides details.

Table 4 also provides intuition that transitory price deviations can be explained by inventories and net order balances. Therefore, we add specialist inventory positions and individual net trades to the transitory prices in Equation (5). $D_{i,t}$ is a dummy variable which takes a value of

[7] We note the compensation is complex. Hendershott and Menkveld (2010) shows that price pressure is a cost to intermediaries. The intermediary earns the half spreads along the way and, in terms of cost, trades off the cost of pressuring price and mean-revert inventory more quickly against the cost of being on costly nonzero inventory.

plus one (+1) if both $Spec_{i,t}$ and $\Delta Indv_{it}$ are positive, negative one (-1) if both variables are negative, and zero (0) otherwise. The dummy variable, $D_{i,t}$, allows us to estimate interaction effects between specialists' and individuals' trades.

5.1 Specialist Trading Variables

We focus on the role of specialists' inventories by restricting $\kappa_i^{indv} = 0$ in Equation (4) and $\alpha_i^{indv} = 0$, and $D_{i,t} = 0$ in Equation (5).

[Insert Table 5]

Table 5 reports our estimates. For both the efficient price equation and the transitory price equation, we report three facets of results. First, we see that κ_i^{spec} = -1.08 and the negative value indicates that specialists face adverse selection worries. Their inventories tend to be high as price are falling and their inventories tend to be low/negative as prices are rising. We can interpret the coefficient as the amount of fundamental price movement associated with every dollar of trading.

Second, to quantify the average effect associated with the adverse selection, we multiply the κ_i^{spec} coefficient by the standard deviation of $\tilde{S}pec_{i,t}$. We see the total effect is 269 bp on average.

Finally, to assess the economic magnitude of a 269 bp move in prices, we compare the number to the 954 bp shown in third column. Specialists' trading can roughly explain $\left(\frac{269^2}{954^2}\right)$ or 8% of the the permanent variance.

The key parameter in Table 5 is α_i^{spec} and its estimated value of -0.25 has the conjectured negative sign. Specialists' inventories are high during times of temporary negative shocks. In other words, specialists absorb excess selling pressure and are partially compensated for providing liquidity via buying at temporarily low prices (with the proviso noted in Footnote 7). The α_i^{spec} coefficient is statistically significant as 263 of the 1,019 stocks produce significantly negative estimates, 683 are insignificant, and only 73 are significantly positive.

The average temporary price pressure is 158 bp as shown in the fifth column. Estimated price pressure varies from 121 bp for large-cap stocks to 248 bp for small-cap stocks. We compare the average value to the transitory price movement of 452 bp shown in Table 3. For all stocks, we estimate that specialists' trading can explain $\left(\frac{158^2}{452^2}\right)$ or 12.22% of temporary price variation.

5.2 Individual Trading Variables

We focus on the role of individuals' net trades by restricting $\kappa_i^{spec} = 0$ in Equation (4) and $\alpha_i^{spec} = 0$, and $D_{i,t} = 0$ in Equation (5).

[Insert Table 6]

Table 6 reports our estimates. We see κ_i^{indv} is negative indicating that individuals' face adverse selection worries. While the slope coefficient is smaller in magnitude that the specialists' coefficient (-0.09 vs. -1.08) the average effect is similar in magnitude. The second column multiplies the slope coefficient by the standard deviation of $\Delta \tilde{I}ndv_{i,t}$ to estimate a 268 bp effect. The 268 bp represents 8.3% of total variance calculated as $\frac{268^2}{929^2}$.

We focus again on the α_i^{indv} parameter, which is equal to -0.06 and has the conjectured negative sign. It is statistically significant as 264 of the 1,019 stocks estimates are significantly negative, 692 are insignificant and only 63 are significantly positive. We interpret the results as coming from imperfect risk sharing among individual investors who find it costly to continuously participate in the market.

The conditional price pressure relating to individual investors (the α_i^{indv} coefficient) varies from -0.17 for small-cap stocks to -0.01 for large-cap stocks. The average price pressure explained by individual investors' trades is 166 bp at a monthly frequency. The economic magnitude is similar to what is associated with specialists. We find individuals' net trades explain $\frac{166^2}{452^2}$ or about 13.49% of transitory price volatility.

5.3 Both Specialist and Individual Variables

Our final analysis includes both specialists' and individuals' trading variables in our state space model. The relevant parts of the full statistical model are shown in Equations (4) and (5). The goal of this section is to test whether one group's trading variables tend to "drive out" the other group's variables. Or, do trading variables from both groups combine to explain stock price volatility.

[Insert Table 7]

Table 7 clearly shows that both groups' trading variables play an important role in our state space model. In the efficient price equation, both κ_i^{spec} and κ_i^{indv} remain negative with values of -0.96 and -0.09. Both groups buy as prices are falling and both groups tend to sell as prices are rising. The permanent volatility explained by specialists is 245 bp while the permanent volatility explained by individuals is 259 bp. These values can be compared to an average total permanent volatility of 930 basis points.

The transitory price equation clarifies the value of including both groups. Both α_i^{spec} and α_i^{indv} are negative. Importantly, the interaction coefficient, α_i^D, is also negative. The interaction coefficient indicates that price pressure is disproportionately large at times that both the specialist and the individual are needed to buffer shocks from other investors (i.e. institutional investors).

We calculate the variance of explained transitory price pressure using the following equation:

$$\Big(\alpha_i^{spec}\sigma(Spec)\Big)^2 + \Big(\alpha_i^{indv}\sigma(\Delta Indv)\Big)^2 + Corr\Big(Spec, \Delta Indv\Big)\cdot\Big(\alpha_i^{spec}\sigma(Spec)\Big)\cdot\Big(\alpha_i^{indv}\sigma(\Delta Indv)\Big)$$

where the correlation coefficient is 0.05 has been previously estimated and shown in Table 4. We find that the average price pressure variance identified using trading variables is more than one-fourth of the base-case transitory price deviations (see Table 3). The result can be calculated as shown below and can be interpreted as the fraction of variance explained:

$$\frac{163^2 + 166^2 + 0.05 \times 163 \times 166}{452^2} \quad = \quad 27.15\%$$

The above represents the final contribution of our paper. We find specialists' and individuals' trading variables are 27.15% the magnitude of transitory variation.

6 Conclusions

We began this paper with the observation that many empirical asset pricing studies use monthly stock prices and returns. To better understand liquidity-related "noise" that may exist in monthly prices, we present a state space (statistical) model in which a stock's observable price is composed of two unobservable components. The first component represents the stock's fundamental value while the second represents transitory price pressure.

Our first question asks: How large are transitory price deviations around fundamental values? This paper provides answers. Using our base case state space model, we estimate that transitory price pressure accounts for more than one quarter of efficient price variation at a monthly frequency.

We use a Kalman filter to decompose observable prices into the two unobservable components. We are able to estimate correlations of efficient price changes, transitory price pressures, and trading variables from two groups on the NYSE who have been shown to trade against prices. We show specialists inventories (and individuals' net trades) are negatively correlation with past/current returns and positively correlated with future returns.

Our final set of empirical results address the question: How can we use trading measures from market participants to help answer and quantify questions about transitory price pressure? Specifically, we add our trading variables to the state space model. Doing so gives a plethora of results including: specialists' inventories and individuals' net trades are associated with 1.63% and 1.66% of monthly volatility, respectively. The results are larger for smaller stocks (2.62% and 1.95%). When the two groups trade in concert, their positions/trades explain 27.15% of transitory variance.

We end by noting that our results are consistent with a model of imperfect risk sharing and costly participation. Market makers have low costs and continuously participate in the market. After building positions, they share/spread risk and unwind positions by trading with individuals (whose participation costs are higher).

References

Cheung, Ying-Wong, and Kon S. Lai. 1995. "Lag orders and critical values of augumented Dickey-Fuller test." *Journal of Business and Economic Statistics* 13(3):277–280.

Durbin, James, and Siem Jan Koopman. 2001. *Time Series Analysis by State Space Model*. Oxford: Oxford University Press.

Grossman, Sanford J., and Merton H. Miller. 1988. "Liquidity and Market Structure." *Journal of Finance* 43(3):617–637.

Hendershott, Terrence, and Albert J. Menkveld. 2010. "Price pressure." Working Paper.

Hendershott, Terrence, and Mark S. Seasholes. 2007. "Market Maker Inventories and Stock Prices." *American Economic Review* 97:210–214.

Ho, Thomas, and Hans R. Stoll. 1981. "Optimal dealer pricing under transactions and return uncertainty." *Journal of Financial Economics* 9(1):47–73.

Kaniel, Ron, Gideon Saar, and Sheridan Titman. 2008. "Individual investor trading and stock returns." *Journal of Finance* 63(1):273–310.

Koopman, Siem Jan, Neil Shephard, and Jurgen A. Doornik. 1999. "Statistical algorithms for models in state space using ssfpack 2.2." *Econometrics Journal* 2:113–166.

Lo, Andrew W., Harry Mamaysky, and Jiang Wang. 2004. "Asset prices and trading volume under fixed transactions costs." *Journal of Political Economy* 112(5):1054–1090.

Madhavan, Ananth, and Seymour Smidt. 1993. "An Analysis of Daily Changes in Specialist Inventories and Quotations." *Journal of Finance* 48:1595–1628.

Menkveld, Albert J., Siem Jan Koopman, and Andre Lucas. 2007. "Modeling Around-the-Clock Price Discovery for Cross-Listed Stocks Using State Space Models." *Journal of Business & Economic Statistics* 25:213–225.

Vayanos, Dimitri, and Jiang Wang. 2009. "Liquidity and Asset Prices: A Unified Framework." Working Paper.

A Proofs for Theoretical Framework

We start with the institutional investors (Group a). Agents choose holdings of risky ($\bar{x}_t^a$) and riskfree assets (B_t^a) to maximize the expected utility of t=3 wealth ($\mathbb{E}\left[U(W_3^a)\right]$) subject to the following constraints:

$$\begin{aligned} W_3^a &= B_2^a + \bar{x}_2^a \tilde{P}_3 + (\tilde{D} - \bar{D}) && (6) \\ W_2^a &= B_2^a + \bar{x}_2^a \tilde{P}_2 = B_1^a + \bar{x}_1^a \tilde{P}_2 && (7) \\ W_1^a &= B_1^a + \bar{x}_1^a \tilde{P}_1 = W_0^a + \theta \tilde{P}_1 && (8) \end{aligned}$$

where θ represents the initial endowment of the risky asset and W_0^a represents other initial wealth. We eliminate B_1^a and B_2^a from the equations above to obtain

$$W_3^a = W_0 + (\tilde{P}_2 - \tilde{P}_1)(\bar{x}_1^a - \theta) + (\tilde{P}_3 - \tilde{P}_2)(\bar{x}_2^a - \theta) + \tilde{P}_3\theta + (\tilde{D} - \bar{D})$$

where $\bar{x}_t^a - \theta$ is the trader's excess demand. Let $x_t^a \equiv \bar{x}_t^a - \theta$, we have

$$W_3^a = W_0 + (\tilde{P}_2 - \tilde{P}_1)x_1^a + (\tilde{P}_3 - \tilde{P}_2)x_2^a + \tilde{P}_3\theta + (\tilde{D} - \bar{D}) \tag{9}$$

We assume that the agents have exponential utility function, i.e., $U(W) = -e^{-\alpha W}$. By backward induction, we solve for the optimal excess demand at t=2

$$\max_{x_2^a} \mathbb{E}_2 \left[U(W_2) - P_2\theta + (\tilde{P}_3 - P_2)x_2^a + \tilde{P}_3\theta + (\tilde{D} - \bar{D}) \right] \tag{10}$$

Using the exponential utility function,

$$\mathbb{E}_2\left[U(W_3^a)\right] = \exp\left\{ -\alpha[W_2 - P_2\theta + (\mathbb{E}_2[\tilde{D}] - P_2)x_2^a + \mathbb{E}_2[\tilde{D}]\theta + \tilde{\epsilon}_1 - \frac{1}{2}\alpha\sigma_2^2(x_2^a + \theta + 1)^2] \right\} \tag{11}$$

The optimal value for x_2^a with all means/variances conditional on the information at t=2:

$$x_2^a = \frac{\mathbb{E}_2[\tilde{D}] - P_2}{\alpha\sigma_2^2} - (\theta + 1). \tag{12}$$

For the individual investors, only fraction $\lambda \in [0, 1]$ of the group participate at t=1 while all participate at t=2. Denote the excess demand of those participating at t=1 as $x_t^b(p)$ and the excess demand of those not participating at t=1 as $x_t^b(np)$. At t=2, the excess demands of risky assets are:

$$x_2^b(p) = x_2^b(np) = \frac{\mathbb{E}_2[\tilde{D}] - P_2}{\alpha\sigma_2^2} - (\theta - 1). \tag{13}$$

Market makers have the same utility function and initial endowment as other investor types, but they don't have a non-tradable endowment of wealth at t=3. Their total excess demand at t=2 is

$$x_2^m = \frac{\mathbb{E}_2[\tilde{D}] - P_2}{\alpha\sigma_2^2} - \theta. \tag{14}$$

The market-clearing condition at t=2 requires that the aggregate excess demand is 0:

$$x_2^a + \lambda x_2^b(p) + (1-\lambda)x_2^b(np) + x_2^m = 0 \tag{15}$$

i.e.,

$$\frac{\mathbb{E}_2[\tilde{D}] - P_2}{\alpha\sigma_2^2} - (\theta+1) + \frac{\mathbb{E}_2[\tilde{D}] - P_2}{\alpha\sigma_2^2} - (\theta-1) + \frac{\mathbb{E}_2[\tilde{D}] - P_2}{\alpha\sigma_2^2} - \theta = 0 \tag{16}$$

Equation (16) implies:

$$\frac{\mathbb{E}_2[\tilde{D}] - P_2}{\alpha\sigma_2^2} = \theta \tag{17}$$

The equilibrium excess demand at t=2 of each group is

$$x_2^a = -1 \tag{18}$$
$$x_2^b(p) = +1 \tag{19}$$
$$x_2^b(np) = +1 \tag{20}$$
$$x_2^m = 0 \tag{21}$$

At t=1, substitute equation (17) and (18) into (9) to get

$$W_3^a = W_0 + \left(\mathbb{E}_2[\tilde{D}] - \theta\alpha\sigma_2^2 - P_1\right) x_1^a + (\tilde{P}_3 - \mathbb{E}_2[\tilde{D}] + \theta\alpha\sigma_2^2)(-1) + \tilde{P}_3\theta + (\tilde{D} - \bar{D}) \tag{22}$$

Solve the maximization problem of Group a and find the excess demand at t=1

$$x_1^a = \frac{\bar{D} - \theta\alpha\sigma_2^2 - P_1}{\alpha\sigma_1^2} - (\theta+1) \tag{23}$$

Similarly, we get the excess demands of the other two groups at t=1

$$x_1^b(p) = \frac{\bar{D} - \theta\alpha\sigma_2^2 - P_1}{\alpha\sigma_1^2} - (\theta-1) \tag{24}$$
$$x_1^b(np) = 0 \tag{25}$$
$$x_1^m = \frac{\bar{D} - \theta\alpha\sigma_2^2 - P_1}{\alpha\sigma_1^2} - \theta \tag{26}$$

Market clearing at t=1 requires

$$x_1^a + \lambda x_1^b(p) + (1-\lambda)x_1^b(np) + x_1^m = 0 \tag{27}$$

Therefore,

$$\frac{\bar{D} - \theta\alpha\sigma_2^2 - P_1}{\alpha\sigma_1^2} = \theta + \frac{1-\lambda}{\lambda+2} \tag{28}$$

and the equilibrium excess demand of the three groups at t=1 is

$$x_1^a = -\frac{2\lambda+1}{\lambda+2} \tag{29}$$

$$x_1^b(p) = \frac{3}{\lambda+2} \tag{30}$$

$$x_1^b(np) = 0 \tag{31}$$

$$x_1^m = \frac{1-\lambda}{\lambda+2} \tag{32}$$

Now we determine the participation intensity λ of individual investors. The advantage of participating at t=1 is the ability to better hedge non-traded risk and to trade at better prices (due to matching institutional demand). However, there is a cost κ for trading at t=1. In equilibrium, individual investors are indifference between participating and not participating at t=1. If individuals participate at both dates, the expected utility of wealth at t=1 is

$$\mathbb{E}_1[U(W_3^b(p))] = \exp\Big\{ -\alpha[W_0 + (\mathbb{E}_1[\tilde{P}_2] - P_1)x_1^b(p) + (\mathbb{E}_1[\tilde{P}_3] - \mathbb{E}_1[\tilde{P}_2])x_2^b(p) + \mathbb{E}_1[\tilde{P}_3]\theta - \kappa$$

$$-\frac{1}{2}\alpha\sigma_1^2(x_1^b(p)+\theta-1)^2 - \frac{1}{2}\alpha\sigma_2^2(x_2^b(p)+\theta-1)^2]\Big\}$$

If individuals only participate at t=2, the expected utility of terminal wealth at t=1 is

$$\mathbb{E}_1\left[U(W_3^b(np))\right] = \exp\left\{-\alpha\left(W_0 + \left(\mathbb{E}_1[\tilde{P}_3] - \mathbb{E}_1[\tilde{P}_2]\right)x_2^b(np) + \mathbb{E}_1[\tilde{P}_3]\theta - \frac{1}{2}\alpha\sigma_1^2(\theta-1)^2 - \frac{1}{2}\alpha\sigma_2^2\left(x_2^b(np)+\theta-1\right)^2\right)\right\}$$

B Likelihood Optimization

This section is referenced from Appendix B of Hendershott and Menkveld (2010). The likelihood of the state space model described by Equations (1), (2), and (3) is optimized in essentially three steps so as to minimize the probability of finding a local maximum. The optimization is implemented in Ox using standard optimization routines. It uses a Kalman filter from `ssfpack` which is an add-on package in Ox—see Koopman, Shephard, and Doornik (1999).

1. An OLS regression of log price difference on contemporaneous and lagged f_t yields starting values for β_i and $\beta_{i,0}, \ldots, \beta_{i,3}$. See Equations (2) and (3). These β estimates are fixed at these values until the final step.

2. The likelihood is calculated using the Kalman filter, see Durbin and Koopman (2001), and optimized numerically using the quasi-Newton method developed by Broyden, Fletcher, Goldfarb, and Shanno. In the optimization all parameters are free except for the βs and $(\sigma(\epsilon), \varphi)$. The latter runs over a nine by nine grid where φ ranges from 0.0 to 0.8 and $\sigma(\epsilon)$ ranges from 0 to a stock-specific upper bound that is calculated assuming that 80% of a stock's unconditional variance is price pressure. The likelihoods are compared across all 81 optimizations and the $(\sigma(\epsilon), \varphi)$ value that yields the highest likelihood is kept as starting value for the final optimization. The rationale for this step is to prevent numerical instability of the quasi-Newton optimization. That is, if all parameters are free on arbitrary starting values the optimization routine often runs off to a persistence parameter φ that approaches its upper bound of and price pressure variance approaches the stock's unconditional variance. The optimizer starts to load the observed price series on two nonstationary series, i.e., the efficient price and the price pressure, and becomes unstable. The Kalman filter is initialized with a diffuse distribution for the unobserved efficient price m_0 and the unconditional price pressure distribution for s_0, i.e., $s_0 \sim \mathrm{N}(0, \frac{\sigma^2(\epsilon)}{1-\varphi^2})$.

3. The likelihood is optimized where all parameters are free and starting values for: β_i, $\beta_{i,0}, \ldots, \beta_{i,3}$, $\sigma(\epsilon)$, φ are equal to those found in steps 1 and 2.

This procedure proves numerically stable as we have strong convergence in the likelihood optimization for all of our stock-year samples, i.e., convergence both in (i) the likelihood elasticity with respect to its parameters and (ii) the one-step change in parameter values. They both become arbitrarily small.

C Efficient Price Estimation and Kalman Filters

Consider the basic state space model (without trading variables) as shown in Equations (1), (2), and (3). The Kalman filter is a recursive algorithm used to evaluate moments of the state vector $\hat{m}_{i,t+1|t}$ conditional on the data set $\mathcal{P}_{i,t} = (p_{i,1}, \ldots, p_{i,t})$, that is:

$$\hat{m}_{i,t+1|t} = E(m_{i,t+1}|\mathcal{P}_{i,t})$$

The estimated state vector $\hat{m}_{i,t+1|t}$ is the predicted state as it is the one-step ahead prediction of state vector. It is obtained from the `ssfmomentest` function in `ssfpack`—see Koopman, Shephard, and Doornik (1999)). In the derivation of Kalman filter(see Durbin and Koopman (2001), we can link the posterior state estimate $\hat{m}_{i,t|t} = E(m_{i,t}|\mathcal{P}_{i,t})$, which is conditional expectation of efficient price $m_{i,t}$ given the observation up to and including time t, to the predicted state. The derivation is the following:

$$\begin{aligned} E(m_{i,t+1}|\mathcal{P}_{i,t}) &= E(m_{i,t} + \beta_i f_{t+1} + w_{i,t+1}|\mathcal{P}_{i,t}) \\ &= E(m_{i,t}|\mathcal{P}_{i,t}) + E(\beta_i f_{t+1}|\mathcal{P}_{i,t}) + E(w_{i,t+1}|P_{i,t}) \\ &= E(m_{i,t}|\mathcal{P}_{i,t}) \end{aligned}$$

Therefore, $\hat{m}_{i,t|t} = \hat{m}_{i,t+1|t}$. We study changes to the estimated efficient price $\Delta\hat{m}_{i,t}$ in the cross-correlation analysis in Table 4.

D Initializing the Kalman Filter

In the state space model, the distribution of the initial state $s_{i,0}$ follows $s_{i,0} \sim N(0, \frac{\sigma_\epsilon^2}{1-\phi^2})$ while the distribution of $m_{i,0}$ is unknown given the state is non-stationary. We follow the convention in Durbin and Koopman (2001) to represent $m_{i,0}$ as having a diffuse prior density, i.e., a random variable with infinite variance. Some people suggest an alternative approach which assumes that $m_{i,0}$ is an unknown constant and estimate it by maximum likelihood from the first observation $p_{i,0}$. In §5.7.3 of Durbin and Koopman (2001), they show that the diffuse initialization of the Kalman filter is the same as assuming that the initial state is fixed and unknown and estimating it from the first observation for the general linear Gaussian state space model. The reason for adopting diffuse initialization is that it is more efficient than the other approach in computation (see §5.7.5 of Durbin and Koopman (2001)).

E Estimating the Magnitude of Transitory Volatility

If the transitory price follows $s_t = \phi s_{t-1} + \epsilon_t$, the first and second order autocorrelation of midquote return are:

$$\rho_1 = \frac{-(1-\phi)\cdot\sigma(\epsilon)^2}{(1+\phi)\cdot\sigma(w)^2 + 2\sigma(\epsilon)^2}$$

$$\rho_2 = \frac{-\phi(1-\phi)\cdot\sigma(\epsilon)^2}{(1+\phi)\cdot\sigma(w)^2 + 2\sigma(\epsilon)^2}$$

The implied ratio of transitory volatility over permanent volatility is:

$$\begin{aligned}
\frac{\frac{\sigma(\epsilon)^2}{1-\phi^2}}{\sigma(w)^2} &= -\frac{\rho_1^3}{(\rho_1 + 2\rho_1^2 - \rho_2)(\rho_1 - \rho_2)} \\
&= -\frac{(-0.06)^3}{(-0.06 + 2(-0.06)^2 + 0.03)(-0.06 + 0.03)} \\
&= 31.58\%
\end{aligned}$$

F Variable Definitions

**	Indicates a variable used throughout the paper.
*	Indicates a variable used infrequently in Tables 2 to 7.

Price Variables:

	$P_{i,t}$	Price of stock i, in dollars, at the end of month t.
	$\overline{P}_i$	Average price of stock i, in dollars, over the sample period.
	$p^{ln}_{i,t}$	Natural log of stock i's price at the end of month t.
**	$p_{i,t}$	Adjusted price of stock i's after subtracting its required return. Defined as: $p_{i,t} = p^{ln}_{i,t} - \delta_{i,t}$ where $\delta_{i,t}$ is defined below.
	$\hat{m}_{i,t}$	Estimate of the efficient price from the Kalman filter: $\hat{m}_{i,t} = E\left[m_{i,t}\vert\mathcal{P}_{i,t}\right]$ where $\mathcal{P}_{i,t}$ is the set containing all current and past prices.
*	$\hat{s}_{i,t}$	Estimate of the transitory price pressure: $\hat{s}_{i,t} = p_{i,t} - \hat{m}_{i,t}$.

Market Capitalizations and Weights:

$MktCap_{i,t}$	Market capitalization of stock i, in dollars, at the end of month t.
$\overline{MktCap}_i$	Average market cap.n of stock i, in dollars, over the sample period.
$\omega_{i,t}$	Weight of stock i in our "market" of 1,019 stocks: $\omega_{i,t} = \frac{\overline{MktCap}_i}{\sum_{i=1}^{N} \overline{MktCap}_i}$.

Return Variables:

	$r_{i,t}$	Return of stock i's over month t: $r_{i,t} = p^{ln}_{i,t} - p^{ln}_{i,t-1}$.
	$r_{f,t}$	Return of riskfree rate over month t and from Ken French's website.
	$r^{std}_{i,t}$	Standardized return of stock i's over month t: $r^{std}_{i,t} = \frac{r_{i,t} - \overline{r_{i,t}}}{std(r_{i,t})}$.
	r_t	Market-wide (return) common factor and equal: $r_t = \sum_i \omega_{i,t} r^{std}_{i,t}$.
**	f_t	Innovation in market-wide returns. Defined as: $f_t = \xi_t$ from the regression: $r_t = \alpha + \phi_1 r_{t-1} + \phi_2 r_{t-2} + \phi_3 r_{t-3} + \phi_4 r_{t-4} + \xi_t$.
*	$r^{idio}_{i,t}$	Idiosyncratic portion of stock i's return. Defined as: $r^{idio}_{i,t} = \xi_{i,t}$ from the regression: $r_{i,t} = \alpha + \phi_0 f_t + \ldots + \phi_4 f_{t-4} + \xi_{i,t}$.
**	$\Delta\hat{m}_{i,t}$	Return of the estimated efficient (unobservable) price. Defined as: $\Delta m_{i,t} = E\left[m_{i,t}\vert\mathcal{P}_{i,t}\right] - E\left[m_{i,t-1}\vert\mathcal{P}_{i,t-1}\right]$.

Required Return Adjustment:

Step 1: Do a fixed-effects panel regression with all 1,019 stocks using the whole sample period: $r_{i,t} = \alpha + \beta_0 f_t + \beta_1 f_{t-1} + \beta_2 f_{t-2} + \beta_3 f_{t-3} + \beta_4 f_{t-4} + \varepsilon_{i,t}$.

Step 2: Do a stock-by-stock regression of the form:
$r_{i,t} = \alpha_i + \beta_{i,0} f_t + \beta_{i,1} f_{t-1} + \beta_{i,2} f_{t-2} + \beta_{i,3} f_{t-3} + \beta_{i,4} f_{t-4} + \varepsilon_{i,t}$.

Step 3: Calculate stock i's beta as: $\beta_i = \frac{\sum_{j=0}^{4} \beta_{i,j}}{\sum_{j=0}^{4} \beta_j}$.

Step 4: Calculate the required return as: $\delta_{i,t} = r_{f,t} + \beta_i \left(1.06^{\frac{1}{12}} - 1\right)$.

Specialists' Inventory Variables:

$Spec^{sh}_{i,t}$ Specialist's inventory (in shares) of stock i at the end of month t.

$Spec^{\$}_{i,t}$ Specialist's inventory (in dollars) of stock i at the end of month t
Defined as: $Spec^{\$}_{i,t} = Spec^{sh}_{i,t} \times \overline{P}_i$.

$Spec^{std}_{i,t}$ Standardized value of specialist's inventory of stock i's at the end of month t. Defined as: $Spec^{std}_{i,t} = \frac{Spec^{\$}_{i,t} - \overline{Spec}^{\$}_{i,t}}{std\left(Spec^{\$}_{i,t}\right)}$.

γ^{Spec}_t Common (market-wide) inventory factor at the end of month t.
Defined as: $\gamma^{Spec}_t = \sum_i \omega_i \times Spec^{std}_{i,t}$.

** $Spec_{i,t}$ Idiosyncratic part of specialist's inventory. Defined as: $Spec_{i,t} = \varepsilon_{i,t}$ from the regression: $Spec^{\$}_{i,t} = \alpha + \beta \cdot \gamma^{Spec}_t + \varepsilon_{i,t}$.

* $\Delta Spec_{i,t}$ Defined as: $Spec_{i,t} - Spec_{i,t-1}$.

* $\tilde{S}pec_{i,t}$ Defined as the residual from an AR(1): $\tilde{S}pec_{i,t} = \varepsilon_{i,t}$ from the regression $Spec_{i,t} = \phi_0 + \phi_1 Spec_{i,t-1} + \varepsilon_{i,t}$.

Individuals' Trading Variables:

	$Indv^{sh}_{i,t}$	Individuals' inventories/holdings (in shares) of stock i at the end of month t. Assumed to be zero at the beginning of the sample period.
	$Indv^{\$}_{i,t}$	Individuals' inventories/holdings (in dollars) of stock i at the end of month t. Defined as: $Indv^{\$}_{i,t} = Indv^{sh}_{i,t} \times \overline{P}_i$.
	$\Delta Indv^{\$}_{i,t}$	Individuals' net trading (in dollars) of stock i's at the end of month t. Defined as: $\Delta Indv^{\$}_{i,t} = Indv^{\$}_{i,t} - Indv^{\$}_{i,t-1}$.
	$Indv^{std}_{i,t}$	Standardized value of Individuals' inventories/holdings of stock i at the end of month t. Defined as: $Indv^{std}_{i,t} = \frac{Indv^{\$}_{i,t} - \overline{Indv}^{\$}_{i,t}}{std\left(Spec^{\$}_{i,t}\right)}$.
	γ^{Indv}_t	Common (market-wide) inventory factor at the end of month t. Defined as: $\gamma^{Indv}_t = \sum_i \omega_i \times Indv^{std}_{i,t}$.
	$\Delta\gamma^{Indv}_t$	Net trading of common factor over month t: $\Delta\gamma^{Indv}_t = \gamma^{Indv}_t - \gamma^{Indv}_{t-1}$.
*	$Indv_{i,t}$	Idiosyncratic part of individuals' inventory. Defined as: $Indv_{i,t} = \varepsilon_{i,t}$ from the regression: $Indv^{\$}_{i,t} = \alpha + \beta \cdot \gamma^{Indv}_t + \varepsilon_{i,t}$.
**	$\Delta Indv_{i,t}$	Idiosyncratic part of net trades. Defined as: $\Delta Indv_{i,t} = \varepsilon_{i,t}$ from the regression $\Delta Indv^{\$}_{i,t} = \alpha + \beta \cdot \Delta\gamma^{Indv}_t + \varepsilon_{i,t}$.
	$\tilde{I}ndv_{i,t}$	Defined as the residual from an AR(1): $\tilde{I}ndv_{i,t} = \varepsilon_{i,t}$ from the regression: $Indv_{i,t} = \phi_0 + \phi_1 Indv_{i,t-1} + \varepsilon_{i,t}$.
*	$\Delta\tilde{I}ndv_{i,t}$	Defined as the residual from an AR(1): $\Delta\tilde{I}ndv_{i,t} = \varepsilon_{i,t}$ from the regression: $\Delta Indv_{i,t} = \phi_0 + \phi_1 \Delta Indv_{i,t-1} + \varepsilon_{i,t}$.

Table 1: Summary Statistics

The table presents summary statistics of our monthly data. Four sources provide data used in this paper: SPETS, CAUD, TAQ, and CRSP. We construct a balanced panel that contains monthly observations of 1,019 NYSE common stocks starting January 1999 and ending December 2005. Stocks are sorted into quintiles based on market capitalization. The first quintile (Q1) contains the smallest stocks. Panel A shows quintile means and the within standard deviation. Panel B considers idiosyncratic variables that are already demeaned. Therefore, we report only within standard deviations.

Panel A: Raw Variables

Variable	Description	Units	Source	Small Q1	Q2	Q3	Q4	Large Q5	Within Stdev[a]
MarCap_{it}	Market capitalization	\$ billion	CRSP	0.26	0.75	1.63	4.19	33.90	6.57
Volume_{it}^{sh}	Average daily share volume	Millions	TAQ	0.05	0.14	0.29	0.69	2.39	0.66
P_{it}	Closing midquote price[b]	\$	NYSE	16.61	25.22	32.94	40.42	58.67	18.74
$Spec_{it}^{sh}$	Specialists' closing inventory	1,000 shares	NYSE	5.93	2.77	3.12	5.40	13.04	40.55
$Spec_{it}^{\$}$	" " "	\$ 1,000[b]	NYSE/CRSP	69.48	50.52	63.02	134.94	530.98	1,315.4
$\Delta Indv_{it}^{sh}$	Individual's net trades	1,000 shares	NYSE	-12.85	-29.66	-44.88	-81.43	-313.32	529.3
$\Delta Indv_{it}^{\$}$	" " "	\$ 1,000[b]	NYSE/CRSP	-195	-683	-1,160	-2,589	-14,116	18,170

Notes: The number of observations is equal to N × T = 1,019 × 84 = 85,586

[a] Based on deviations from time-series means i.e. $x_{i,t}^{*} = x_{i,t} - \overline{x}_i$.

[b] We adjust all price series to account for stock splits and dividends.

Panel B: Stdev of Idiosyncratic Variables

Variable	Description	Units		Small Q1	Q2	Q3	Q4	Large Q5	Within Stdev
$Spec_{i,t}$	Specialists' inventory	$ 1,000	Mean-Avg	0.0	0.0	0.0	0.0	0.0	0.0
$Spec_{i,t}$	" "	″	Std-Avg	275.4	347.4	553.8	888.8	2,623.4	1,275.0
$\Delta Indv_{i,t}$	Individual's net trades	$ 1,000	Mean-Avg	0.0	0.0	0.0	0.0	0.0	0.0
$\Delta Indv_{i,t}$	" "	″	Std-Avg	1,088.0	2,883.8	3,978.5	7,574.8	39,505.0	18,062.0
$r_{i,t}^{idio}$	Idiosyncratic returns	%	Mean-Avg	0.0	0.0	0.0	0.0	0.0	0.0
$r_{i,t}^{idio}$	" "	″	Stc-Avg	13.74	12.23	11.62	11.22	10.31	11.88

Table 2: Augmented Dickey-Fuller Tests

This table presents results of augmented Dickey-Fuller tests. We report the cross-sectional mean of the β coefficient. We also report the cross-sectional mean of t-statistic associated with the β coefficient. We consider two inventory variables: $Spec_{i,t}$ and $Indv_{i,t}$.

$$\Delta Spec_{i,t} \quad = \quad \alpha + \beta Spec_{i,t-1} + \phi_1 \Delta Spec_{i,t-1} + \ldots + \phi_4 \Delta Spec_{i,t-4} + \varepsilon_{i,t}$$

The p-values, reported in brackets, are based on a test statistic that counts the number of significant augmented Dickey-Fuller test statistic across all stocks estimates in the bin. The test statistic is binomial distributed under the null (we use the 0.10 quantiles in the DF-test). The data are monthly starting January 1999 and ending December 2005.

		Specialists' Inventories ($Spec_{i,t}$)	Individuals' Net Trades ($Indv_{i,t}$)
All	β-Avg	-0.782	-0.019
	T-Avg	-3.5	-0.82
	P-value	[0.000]	[1.000]
Q1 (Small)	β-Avg	-0.606	-0.023
	T-Avg	-3.3	-0.95
	P-value	[0.000]	[0.953]
Q2	β-Avg	-0.815	-0.020
	T-Avg	-3.6	-0.84
	P-value	[0.000]	[0.999]
Q3	β-Avg	-0.830	-0.018
	T-Avg	-3.6	-0.83
	P-value	[0.000]	[0.999]
Q4	β-Avg	-0.846	-0.020
	T-Avg	-3.6	-0.86
	P-value	[0.000]	[0.997]
Q5 (Large)	β-Avg	-0.813	-0.014
	T-Avg	-3.4	-0.62
	P-value	[0.000]	[0.999]

Table 3: Base Case State Space Model

This table reports estimates from the base case state space model. The first equation shows the observable log price (p_{it}). The second equation show the unobserved efficient price ($m_{i,t}$). The third equation shows the the unobserved transitory price deviations ($s_{i,t}$).

$$\begin{aligned} p_{i,t} &= m_{i,t} + s_{i,t} \\ m_{i,t} &= m_{i,t-1} + \beta_i f_t + w_{i,t} \\ s_{i,t} &= \phi_i s_{i,t-1} + \beta_{i,0} f_t + \ldots + \beta_{i,3} f_{t-3} + \epsilon_{i,t} \end{aligned}$$

Data are monthly starting January 1999 and ending December 2005. The table reports p-values in parentheses. These p-values are based on a test statistic that counts the number of significant t-values across all stocks ML estimates in the bin. The test statistic is binomially distributed under the null (in the t-test, we use the 0.10 quantiles if cross-sectional mean is negative and 0.90 quantiles if cross-sectional mean is positive).

	$\sigma(w)$	ϕ_i	$\sigma(\epsilon)$	$\left(\frac{\sigma(\epsilon)^2}{1-\phi_i^2}\right)^{\frac{1}{2}}$
All	849	0.35	324	452
	(0.000)	(0.000)	(0.000)	(0.000)
Q1 (Small)	1,059	0.33	365	491
	(0.000)	(0.000)	(0.000)	(0.000)
Q2	913	0.35	345	508
	(0.000)	(0.000)	(0.000)	(0.000)
Q3	790	0.34	348	466
	(0.000)	(0.000)	(0.000)	(0.000)
Q4	788	0.41	275	407
	(0.000)	(0.000)	(0.000)	(0.000)
Q5 (Large)	691	0.34	288	389
	(0.000)	(0.000)	(0.000)	(0.000)

Table 4: Correlations of Trading and Price Variables

This table presents correlations of idiosyncratic inventory of NYSE specialists ($Spec_{i,t}$), individuals net trades ($\Delta Indv_{i,t}$), the idiosyncratic part of monthly returns ($r^{idio}_{i,t}$), estimated price pressure ($\hat{s}_{i,t}$), and the return/changes to the estimated efficient price ($\Delta\hat{m}_{i,t}$). A full description of all variables is given in the text and in Appendix F. Efficient prices and price pressure are estimated by a Kalman filter using the statistical base-case model shown in the text. Correlations are calculated on a stock-by-stock basis using the entire January 1999 to December 2005 sample period. The table shows average matrices (across stocks). The Lag 0 matrix shows contemporaneous correlations. The " – " indicates correlation and autocorrelation estimates are not applicable to this table as the underlying variables are themselves model-implied estimates. p-values are reported in brackets and are based on a test statistic that counts the number of significant t-values across all stocks estimates in the bin. The test statistic is binomially distributed under the null (in the t-test, we use the 0.10 quantiles if cross-sectional mean is negative and 0.90 quantiles if cross-sectional mean is positive).

		Lag 0					Lag 1					Lag 2				
		$Spec$	$\Delta Indv$	r^{idio}	$\hat{s}$	$\Delta\hat{m}$	$Spec_{-1}$	$\Delta Indv_{-1}$	r^{idio}_{-1}	$\hat{s}_{-1}$	$\Delta\hat{m}_{-1}$	$Spec_{-2}$	$\Delta Indv_{-2}$	r^{idio}_{-2}	$\hat{s}_{-2}$	$\Delta\hat{m}_{-2}$
All	$Spec$	1.00	0.05 (0.000)	-0.22 (0.000)	-0.09 (0.000)	-0.23 (0.000)	0.15 (0.000)	0.01 (0.000)	-0.04 (0.000)	0.01 (0.000)	-0.05 (0.000)	0.11 (0.000)	0.01 (0.000)	-0.03 (0.000)	0.02 (0.000)	-0.04 (0.000)
	$\Delta Indv$		1.00	-0.17 (0.000)	-0.12 (0.000)	-0.17 (0.000)	0.07 (0.000)	0.31 (0.000)	-0.17 (0.000)	-0.06 (0.000)	-0.18 (0.000)	0.03 (0.000)	0.16 (0.000)	-0.08 (0.000)	0.01 (0.000)	-0.09 (0.000)
	r^{idio}			1.00	0.18 (0.000)	0.83 (0.000)	0.05 (0.000)	0.03 (0.000)	-0.06 (0.000)	-0.06 (0.000)	-0.06 (0.000)	0.01 (0.000)	0.02 (0.000)	-0.03 (0.000)	-0.03 (0.243)	-0.03 (0.000)
	$\hat{s}$				1.00	–	-0.06 (0.000)	-0.08 (0.000)	0.07 (0.000)	–	–	-0.05 (0.000)	-0.04 (0.000)	0.04 (0.000)	–	–
	$\Delta\hat{m}$					1.00	0.04 (0.000)	0.03 (0.000)	0.01 (0.710)	–	–	0.01 (0.000)	0.02 (0.000)	-0.00 (0.044)	–	–
Q1	$Spec$	1.00	0.04 (0.000)	-0.17 (0.000)	-0.07 (0.000)	-0.18 (0.000)	0.11 (0.000)	-0.00 (0.030)	-0.03 (0.003)	0.00 (0.109)	-0.03 (0.000)	0.09 (0.000)	-0.01 (0.006)	-0.02 (0.019)	0.01 (0.370)	-0.03 (0.000)
	$\Delta Indv$		1.00	-0.22 (0.000)	-0.12 (0.000)	-0.22 (0.000)	0.05 (0.000)	0.36 (0.000)	-0.16 (0.000)	-0.03 (0.001)	-0.16 (0.000)	-0.00 (0.019)	0.19 (0.000)	-0.04 (0.000)	0.01 (0.000)	-0.05 (0.000)
	r^{idio}			1.00	0.18 (0.000)	0.79 (0.000)	0.04 (0.000)	0.03 (0.048)	-0.07 (0.000)	-0.05 (0.011)	-0.07 (0.000)	0.01 (0.156)	0.02 (0.019)	-0.02 (0.003)	-0.03 (0.553)	-0.03 (0.006)
	$\hat{s}$				1.00	–	-0.04 (0.001)	-0.07 (0.000)	0.05 (0.000)	–	–	-0.03 (0.048)	-0.01 (0.002)	0.02 (0.011)	–	–
	$\Delta\hat{m}$					1.00	0.04 (0.000)	0.03 (0.000)	0.01 (0.370)	–	–	0.01 (0.074)	0.02 (0.006)	-0.00 (0.287)	–	–

<continued on next page>

<continued from previous page>

		Lag 0					Lag 1					Lag 2				
		$Spec$	$\Delta Indv$	r^{idio}	$\hat{s}$	$\Delta\hat{m}$	$Spec_{-1}$	$\Delta Indv_{-1}$	r^{idio}_{-1}	$\hat{s}_{-1}$	$\Delta\hat{m}_{-1}$	$Spec_{-2}$	$\Delta Indv_{-2}$	r^{idio}_{-2}	$\hat{s}_{-2}$	$\Delta\hat{m}_{-2}$
Q2	$Spec$	1.00	0.05	-0.20	-0.08	-0.21	0.11	0.01	-0.04	0.02	-0.05	0.08	-0.00	-0.01	0.02	-0.03
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.081)	(0.000)	(0.034)	(0.000)	(0.000)	(0.304)	(0.012)	(0.081)	(0.001)
	$\Delta Indv$		1.00	-0.19	-0.10	-0.19	0.07	0.32	-0.18	-0.04	-0.19	0.02	0.17	-0.08	0.02	-0.09
				(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	r^{idio}			1.00	0.15	0.81	0.04	0.02	-0.04	-0.06	-0.05	0.01	0.01	-0.02	-0.03	-0.03
					(0.000)	(0.000)	(0.002)	(0.054)	(0.000)	(0.081)	(0.000)	(0.001)	(0.304)	(0.002)	(0.230)	(0.007)
	$\hat{s}$				1.00	–	-0.05	-0.07	0.06	–	–	-0.03	-0.04	0.02	–	–
							(0.000)	(0.000)	(0.000)			(0.034)	(0.000)	(0.007)		
	$\Delta\hat{m}$					1.00	0.03	0.02	0.01	–	–	0.01	0.02	0.01	–	–
							(0.021)	(0.000)	(0.388)			(0.012)	(0.004)	(0.034)		
Q3	$Spec$	1.00	0.03	-0.19	-0.09	-0.21	0.12	0.01	-0.02	0.02	-0.03	0.06	0.00	-0.02	0.02	-0.03
			(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.081)	(0.001)	(0.001)	(0.000)	(0.081)	(0.000)	(0.034)	(0.000)
	$\Delta Indv$		1.00	-0.16	-0.13	-0.16	0.05	0.28	-0.16	-0.06	-0.17	0.01	0.13	-0.07	0.01	-0.09
				(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.034)	(0.000)
	r^{idio}			1.00	0.19	0.81	0.06	0.03	-0.06	-0.07	-0.07	0.03	0.03	-0.04	-0.03	-0.05
					(0.000)	(0.000)	(0.000)	(0.388)	(0.000)	(0.004)	(0.000)	(0.000)	(0.007)	(0.000)	(0.230)	(0.000)
	$\hat{s}$				1.00	–	-0.06	-0.08	0.09	–	–	-0.04	-0.04	0.04	–	–
							(0.000)	(0.000)	(0.000)			(0.001)	(0.000)	(0.000)		
	$\Delta\hat{m}$					1.00	0.05	0.03	0.00	–	–	0.03	0.02	-0.02	–	–
							(0.000)	(0.817)	(0.817)			(0.002)	(0.054)	(0.001)		
Q4	$Spec$	1.00	0.06	-0.25	-0.11	-0.26	0.12	0.02	-0.05	-0.00	-0.05	0.09	0.02	-0.04	0.01	-0.05
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.013)	(0.000)	(0.056)	(0.000)	(0.000)	(0.013)	(0.000)	(0.175)	(0.000)
	$\Delta Indv$		1.00	-0.17	-0.12	-0.16	0.08	0.28	-0.17	-0.06	-0.19	0.02	0.15	-0.09	0.01	-0.11
				(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
	r^{idio}			1.00	0.19	0.84	0.06	0.04	-0.04	-0.08	-0.05	0.01	0.01	-0.01	-0.04	-0.02
					(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.085)	(0.124)	(0.022)	(0.085)	(0.085)
	$\hat{s}$				1.00	–	-0.05	-0.08	0.07	–	–	-0.05	-0.05	0.05	–	–
							(0.000)	(0.000)	(0.000)			(0.000)	(0.000)	(0.000)		
	$\Delta\hat{m}$					1.00	0.04	0.04	0.02	–	–	0.01	0.03	0.01	–	–
							(0.000)	(0.000)	(0.175)			(0.013)	(0.004)	(0.237)		
Q5	$Spec$	1.00	0.07	-0.30	-0.11	-0.30	0.30	0.03	-0.08	0.01	-0.09	0.21	0.02	-0.06	0.03	-0.08
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.021)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	$\Delta Indv$		1.00	-0.13	-0.13	-0.12	0.12	0.32	-0.19	-0.09	-0.19	0.08	0.18	-0.12	-0.01	-0.13
				(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.007)	(0.000)
	r^{idio}			1.00	0.17	0.88	0.04	0.05	-0.06	-0.06	-0.06	-0.00	0.02	-0.02	-0.02	-0.02
					(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.054)	(0.000)	(0.230)	(0.000)	(0.012)	(0.662)	(0.021)
	$\hat{s}$				1.00	–	-0.09	-0.10	0.07	–	–	-0.08	-0.07	0.05	–	–
							(0.000)	(0.000)	(0.000)			(0.000)	(0.000)	(0.000)		
	$\Delta\hat{m}$					1.00	0.03	0.05	0.01	–	–	-0.00	0.03	-0.01	–	–
							(0.000)	(0.000)	(0.994)			(0.304)	(0.000)	(0.119)		

Table 5: State Space Model with NYSE Specialists' Inventories

This table presents estimates from a state space model that includes NYSE specialists' inventories and is shown below. $p_{i,t}$ is the observable log price of stock i at the end of month t. $m_{i,t}$ is the unobservable efficient price. $s_{i,t}$ is the unobservable transitory price deviation.

$$\begin{aligned} p_{i,t} &= m_{i,t} + s_{i,t} \\ m_{i,t} &= m_{i,t-1} + \beta_i f_t + w_{i,t} \\ w_{i,t} &= \kappa_i^{spec} \tilde{S}pec_{i,t} + u_{i,t} \\ s_{i,t} &= \alpha_i^{spec} Spec_{i,t} + \beta_{i,0} f_t + \ldots + \beta_{i,3} f_{t-3} + \epsilon_{i,t} \end{aligned}$$

Full descriptions and definitions of variables are given in Appendix F. The model is estimated on a stock-by-stock basis using maximum likelihood estimates where the error terms w_{it} and $\epsilon_{i,t}$ are assumed to be normally and independently distributed. The optimization is implemented in Ox with `ssfpack` routines. A Kalman filter is used to evaluate the likelihood function. The table reports p-values in brackets. These values are based on a test statistic that counts the number of significant t-values across all stocks. The test statistic is binomially distributed under the null. In the t-test, we use the 10% percentile if the cross-sectional mean is negative and the 90% percentile if the cross-sectional mean is positive).

	Efficient Price Equation			**Transitory Price Equation**		
	κ_i^{spec}	$\lvert\kappa_i^{spec}\rvert \cdot \sigma(\tilde{S}pec)$	$\sigma(w)$	α_i^{spec}	$\lvert\alpha_i^{spec}\rvert \cdot \sigma(Spec)$	$\sigma(\epsilon)$
All	-1.08	269	954	-0.25	158	159
	(0.000)			(0.000)		
Q1 (Small)	-3.16	441	1,171	-0.50	248	181
	(0.000)			(0.000)		
Q2	-1.08	281	1,025	-0.38	170	155
	(0.000)			(0.000)		
Q3	-0.53	212	909	-0.24	125	158
	(0.000)			(0.000)		
Q4	-0.44	226	893	-0.09	124	133
	(0.000)			(0.000)		
Q5 (Large)	-0.16	183	769	-0.05	121	169
	(0.000)			(0.000)		

Table 6: State Space Model with Individuals' Net Trades

This table presents estimates from a state space model that includes individuals' net trades and is shown below. $p_{i,t}$ is the observable log price of stock i at the end of month t. $m_{i,t}$ is the unobservable efficient price. $s_{i,t}$ is the unobservable transitory price deviation.

$$\begin{aligned} p_{i,t} &= m_{i,t} + s_{i,t} \\ m_{i,t} &= m_{i,t-1} + \beta_i f_t + w_{i,t} \\ w_{i,t} &= \kappa_i^{indv} \Delta \tilde{I}ndv_{i,t} + u_{i,t} \\ s_{i,t} &= \alpha_i^{indv} \Delta Indv_{i,t} + \beta_{i,0} f_t + \ldots + \beta_{i,3} f_{t-3} + \epsilon_{i,t} \end{aligned}$$

Full descriptions and definitions of variables are given in Appendix F. The model is estimated on a stock-by-stock basis using maximum likelihood estimates where the error terms $w_{i,t}$ and $\epsilon_{i,t}$ are assumed to be normally and independently distributed. The optimization is implemented in Ox with `ssfpack` routines. A Kalman filter is used to evaluate the likelihood function. The table reports p-values in brackets. These values are based on a test statistic that counts the number of significant t-values across all stocks. The test statistic is binomially distributed under the null. In the t-test, we use the 10% percentile if the cross-sectional mean is negative and the 90% percentile if the cross-sectional mean is positive).

	Efficient Price Equation			**Transitory Price Equation**		
	κ_i^{indv}	$\lvert\kappa_i^{indv}\rvert \cdot \sigma(\Delta \tilde{I}ndv)$	$\sigma(w)$	α_i^{indv}	$\lvert\alpha_i^{indv}\rvert \cdot \sigma(\Delta Indv)$	$\sigma(\epsilon)$
All	-0.09	268	929	-0.06	166	227
	(0.000)			(0.000)		
Q1 (Small)	-0.22	267	1,123	-0.17	197	269
	(0.000)			(0.000)		
Q2	-0.11	268	1,000	-0.06	176	241
	(0.000)			(0.000)		
Q3	-0.06	252	878	-0.03	133	234
	(0.000)			(0.000)		
Q4	-0.04	282	880	-0.01	157	187
	(0.000)			(0.000)		
Q5 (Large)	-0.01	271	762	-0.01	164	203
	(0.000)			(0.000)		

Table 7: State Space Model with Both Specialists and Individuals

This table presents estimates from a state space model that includes NYSE specialists' inventories and individuals' net trading. $p_{i,t}$ is the observable log price of stock i at the end of month t. $m_{i,t}$ is the unobservable efficient price. $s_{i,t}$ is the unobservable transitory price deviation.

$$\begin{aligned}
p_{i,t} &= m_{i,t} + s_{i,t} \\
m_{i,t} &= m_{i,t-1} + \beta_i f_t + w_{i,t} \\
w_{i,t} &= \kappa_i^{spec} \tilde{S}pec_{i,t} + \kappa_i^{indv} \Delta \tilde{I}ndv_{i,t} + u_{i,t} \\
s_{i,t} &= \alpha_i^{spec} Spec_{i,t} + \alpha_i^{indv} \Delta Indv_{i,t} + \alpha_i^D D_{i,t} + \beta_{i,0} f_t + \ldots + \beta_{i,3} f_{t-3} + \epsilon_{i,t}
\end{aligned}$$

Full descriptions and definitions of variables are given in Appendix F. The model is estimated on a stock-by-stock basis using maximum likelihood estimates where the error terms w_{it} and $\epsilon_{i,t}$ are assumed to be normally and independently distributed. The optimization is implemented in Ox with `ssfpack` routines. A Kalman filter is used to evaluate the likelihood function. The table reports p-values in brackets. These values are based on a test statistic that counts the number of significant t-values across all stocks. The test statistic is binomially distributed under the null. In the t-test, we use the 10% percentile if the cross-sectional mean is negative and the 90% percentile if the cross-sectional mean is positive).

	Efficient Price Equation					Transitory Price Equation					
	κ_i^{spec}	$\lvert\kappa_i^{spec}\rvert\times$ $\sigma(\tilde{S}pec)$	κ_i^{indv}	$\lvert\kappa_i^{indv}\rvert\times$ $\sigma(\Delta\tilde{I}ndv)$	$\sigma(w)$	α_i^{spec}	$\lvert\alpha_i^{spec}\rvert\times$ $\sigma(Spec)$	α_i^{indv}	$\lvert\alpha_i^{indv}\rvert\times$ $\sigma(\Delta Indv)$	α_i^D	$\sigma(\epsilon)$
All	-0.96	245	-0.09	259	930	-0.26	163	-0.04	166	-57.52	185
	(0.000)		(0.000)			(0.000)		(0.000)		(0.000)	
Q1 (Small)	-2.84	401	-0.23	251	1,131	-0.63	262	-0.14	195	-74.19	208
	(0.000)		(0.000)			(0.000)		(0.000)		(0.000)	
Q2	-0.95	262	-0.11	258	998	-0.30	167	-0.04	176	-86.84	192
	(0.000)		(0.000)			(0.000)		(0.000)		(0.000)	
Q3	-0.49	200	-0.07	253	879	-0.20	129	-0.02	139	-63.63	192
	(0.000)		(0.000)			(0.000)		(0.000)		(0.000)	
Q4	-0.38	203	-0.03	273	881	-0.11	129	-0.01	157	-26.78	153
	(0.000)		(0.000)			(0.000)		(0.012)		(0.021)	
Q5 (Large)	-0.13	160	-0.01	262	761	-0.05	126	-0.00	162	-35.82	179
	(0.000)		(0.000)			(0.000)		(0.000)		(0.002)	