No. 2010/14

Price Pressures

Terrence Hendershott and Albert J. Menkveld

Center for Financial Studies
Goethe-Universität Frankfurt ▪ House of Finance
Grüneburgplatz 1 ▪ 60323 Frankfurt ▪ Deutschland

Telefon: +49 (0)69 798-30050
Fax: +49 (0)69 798-30077
http://www.ifk-cfs.de ▪ E-Mail: info@ifk-cfs.de

Center for Financial Studies

The *Center for Financial Studies* is a nonprofit research organization, supported by an association of more than 120 banks, insurance companies, industrial corporations and public institutions. Established in 1968 and closely affiliated with the University of Frankfurt, it provides a strong link between the financial community and academia.

The CFS Working Paper Series presents the result of scientific research on selected topics in the field of money, banking and finance. The authors were either participants in the Center´s Research Fellow Program or members of one of the Center´s Research Projects.

If you would like to know more about the *Center for Financial Studies*, please let us know of your interest.

Prof. Michalis Haliassos, Ph.D. Prof. Dr. Jan Pieter Krahnen Prof. Dr. Uwe Walz

Center for Financial Studies
Goethe-Universität ■ House of Finance
Grüneburgplatz 1 ■ 60323 Frankfurt am Main ■ Deutschland

Telefon: +49 (0)69 798-30050
Fax: +49 (0)69 798-30077
http://www.ifk-cfs.de ■ E-Mail: info@ifk-cfs.de

CFS Working Paper No. 2010/14

Price Pressures*

Terrence Hendershott[1] and Albert J. Menkveld[2]

February 10, 2010

Abstract:
We study price pressures in stock prices—price deviations from fundamental value due to a risk-averse intermediary supplying liquidity to asynchronously arriving investors. Empirically, twelve years of daily New York Stock Exchange intermediary data reveal economically large price pressures. A $100,000 inventory shock causes an average price pressure of 0.28% with a half-life of 0.92 days. Price pressure causes average transitory volatility in daily stock returns of 0.49%. Price pressure effects are substantially larger with longer durations in smaller stocks. Theoretically, in a simple dynamic inventory model the 'representative' intermediary uses price pressure to control risk through inventory mean reversion. She trades off the revenue loss due to price pressure against the price risk associated with remaining in a nonzero inventory state. The model's closed-form solution identifies the intermediary's relative risk aversion and the distribution of investors' private values for trading from the observed time series patterns. These allow us to estimate the social costs—deviations from constrained Pareto efficiency—due to price pressure which average 0.35 basis points of the value traded.

JEL Classification: G12, G14, D53, D61

Keywords: Liquidity, Inventory Risk, Intermediary, Volatility

* We benefit from in-depth discussions with Viral Acharya, Bruno Biais, Sabrina Buti, Robert Engle, Charles Jones, Jeremy Large, Sunny Li, Virgiliu Midrigan, Mark Seasholes, Stijn van Nieuwerburgh, Matthew Pritsker, Jan van Schuppen, and Pierre-Olivier Weill and audience participation at the Bank of Canada, Cass, Dauphine, ESSEC, the NBER, and the U of Toronto. The project was inspired by the Ljungqvist and Sargent NYU lecture series based on their book Recursive Macroeconomic Theory. Menkveld gratefully acknowledges Robert Engle for sponsoring his NYU visit in 08/10 and VU University Amsterdam for a VU talent grant.

1 Terrence Hendershott, Haas School of Business, University of California at Berkeley, 545 Student Services Bldg #1900, Berkeley, CA 94720, tel +1 510 643 0619, fax +1 510 643 1412, E-mail: hender@haas.berkeley.edu.

2 Albert J. Menkveld, VU University Amsterdam, FEWEB, De Boelelaan 1105, 1081 HV, Amsterdam, Netherlands, tel +31 20 598 6130, fax +31 20 598 6020, E-mail: albertjmenkveld@gmail.com.

The ability and cost to trade large quantities quickly, often referred to as liquidity, plays a fundamental role in facilitating risk-sharing and allocating resources in asset markets. For example, Pastor and Stambaugh (2003) show that liquidity is a priced factor in stock prices and Gavazza (2009) shows that liquidity affects asset prices and utilization in the aircraft market. Furthermore, Levine and Zervos (1998) show that more liquid financial markets increase investment, productivity, and economic growth. Illiquidity can arise for both informational and non-information reasons. Investor demand to trade quickly may move prices permanently due to information, e.g., Evans and Lyons (2002), and temporarily for other reasons, e.g., Kraus and Stoll (1972). Investors' trading demands temporarily distorting prices away from fundamental values, often referred to as price pressure, has been established in specific instances, but systematic evidence has been elusive.

We use 12 years of daily position data from intermediaries on the New York Stock Exchange (NYSE) to measure the amount of liquidity supplied and to characterize the magnitude of price pressures in terms of size, frequency, and duration; these determine the costs investors face when trading large quantities and the effect of investor liquidity demand on the efficiency of prices. We also construct a model that identifies the intermediary's risk aversion and investors' private value distribution from the observed time series patterns of price pressures. These enable decomposition of trading costs and estimation of the social costs (deviations from constrained Pareto efficiency) due to price pressure.

We follow Campbell, Grossman, and Wang (1993) and Pastor and Stambaugh (2003) to focus on the aspect of liquidity associated with temporary price changes induced by order flow (investor net buying or selling). These price pressures arise from shocks to agents' consumption, investment, labor, and other opportunities resulting in liquidity or hedging needs, as in Grossman and Miller (1988). Asynchronous stochastic arrivals by impatient agents unwilling to bear the costs of constantly monitoring the market naturally leads to an intermediary who stands ready to both buy and sell (Townsend (1978)). The intermediary offers buyers and sellers quick exchange (Demsetz (1968)) and temporarily leans against the wind to match asynchronously arriving buyers and sellers across time (Weill (2007)).[1] Stoll (1978) and Grossman and Miller (1988) emphasize that price pressure differs from the bid-ask spread, which measures the intermediary's return if she simultaneously

[1] Duffie, Gârleanu, and Pedersen (2005) and others model search and trading in decentralized over the counter markets.

crosses (executes on both sides of) the trade, one at the bid and the other at the ask.[2]

Price pressures equilibrate liquidity demand and liquidity supply by compensating liquidity providers for bearing risk while supplying immediacy to liquidity demanders. Thus, measures of both liquidity demand and liquidity supply are suitable for characterizing price pressure. Empirical evidence from situations with large investor liquidity demands demonstrates that price pressures can be large.[3] However, long time series of aggregate investor liquidity demand are not available. Prior empirical work on liquidity supply using intermediaries' inventory positions finds support for risk management via inventory control, but weak support for inventories causing price pressure.[4]

To estimate price pressures arising from liquidity supply we employ intermediary data from the New York Stock Exchange from 1994-2005. Market makers who act as intermediaries to supply liquidity on the NYSE, called specialists, are required to report their positions in every security every day. To link these inventory positions with price pressures we use the Kalman filter to estimate a state space model that decomposes stock prices into their fundamental value and 'noise': the random walk component and the stationary component. The stationary component represents pricing error around the fundamental value. The estimation allows the NYSE intermediaries' inventory level to enter directly into the price equation via the stationary noise component. The coefficient on inventory represents the transitory impact of a dollar of intermediary inventory, which we refer to as the conditional price pressure. The standard deviation of the inventory characterizes the frequency of price pressure. Combining the conditional price pressure with its frequency and duration yields the average price pressure which we refer to as the transitory volatility due to price

[2]Grossman and Miller (1988, p.630) also criticize another empirical measure of the liquidity, the liquidity ratio, which is defined as the ratio of average dollar volume of trading to the average price change during some interval. The reciprocal of this, often referred to as the illiquidity ratio, is subject to their same critiques.

[3]Kraus and Stoll (1972) provide some of the first evidence on liquidity demands from block trades causing price pressure. Harris and Gurel (1986) and subsequent papers on additions to the S&P 500 index find evidence for price pressure. Greenwood (2005) extends this with an examination of transitory price effects upon a weighting change to the Nikkei 225. Coval and Stafford (2007) examine price pressures due to mutual fund redemptions. Less directly Campbell, Grossman, and Wang (1993) examine how price pressures leads to stocks' return autocorrelations declining with trading volume. Gabaix, Gopikrishnan, Plerou, and Stanley (2006) examine how price pressures by institutional investors can theoretically affect stock market volatility. Brunnermeier and Pedersen (2008) provide a theoretical model where price pressures affect intermediary capital positions due margin requirements. This can require the intermediary to reduce leverage causing feedback effects in price pressures.

[4]Madhavan and Smidt (1991), Madhavan and Smidt (1993), and Hasbrouck and Sofianos (1993) employ NYSE intermediary inventory data and find evidence supporting risk management, but not price pressure. Hendershott and Seasholes (2007) use a long time series of NYSE data to find cross-sectional evidence of both inventory control and price pressure: a portfolio of stocks where the intermediary is long outperforms a portfolio of stocks where the intermediary is short by 45.4 basis points over two weeks. Our findings extend this portfolio approach to determine the price pressure per dollar of inventory, the average impact of price pressure on stocks' volatility, and the social costs of price pressure.

pressure.

The conditional price pressure for the largest-quintile stocks is 0.02 basis points per one thousand dollars of intermediary inventory.[5] For the smallest-quintile stocks one thousand dollars of inventory results in 1.01 basis points of price pressure. The small conditional price pressure in large stocks is associated with greater frequency of price pressure: the standard deviation of inventory is $1.1 million for the largest stocks versus $165 thousand for the smallest stocks. Inventory positions causing price pressures last longer in small stocks: the half-life of intermediary inventory is 0.55 days for the largest stocks and 2.11 days for the smallest stocks. Combining the conditional size, frequency, and duration of price pressures produces estimates of the daily transitory volatility due to price pressure that range from 0.17% for the largest stocks to 1.20% for smallest stocks. Price pressure contributes substantially to daily volatility in stock prices in small stocks as the average ratio of transitory volatility due to price pressure to permanent volatility is greater than one.

To further understand price pressures we construct a single-asset theoretical model of liquidity supply with stochastically arriving investors with less than perfectly elastic demands to trade. In the infinite-horizon recursive model the intermediary dynamically chooses the prices at which she is willing to buy and sell, the bid and ask prices, respectively. When the intermediary is at her desired inventory position in a security the bid and ask prices symmetrically straddle the security's fundamental value. If a seller then arrives the intermediary buys and her position, also referred to as her inventory level, is higher than desired, exposing her to idiosyncratic price risk. To mitigate this risk the intermediary then stochastically mean reverts her inventory by applying price pressure: she adjusts both the bid and ask prices downward to induce more investor buying selling. In doing this the intermediary bears the cost of setting the average of the bid-ask quotes, the midquote price, to be below the fundamental value. The size of the deviation of the midquote price from the fundamental value is our theoretical and empirical measure of price pressure.

The main innovation of our theoretical approach is to facilitate estimation of the risk aversion of the intermediary and social welfare lost due to deviations from constrained Pareto efficiency. The model together with the time series properties of price pressure identifies low risk aversion for the intermediaries, a 0.10 coefficient of relative risk aversion, and deviations from Pareto efficiency of 0.35 basis points of the value traded. The low risk aversion could arise from risk tolerant capital

[5]Throughout we focus on the idiosyncratic inventory positions as hedging of the systematic inventory risk can be accomplished relatively inexpensively via futures and exchange traded funds based on stock market indices.

migrating to the intermediation sector or be a sign of an agency conflict between firm and the individual traders it employs to act as intermediaries. If the traders have limited liability, then they may be willing to take more risky bets than the firm's owners would prefer.

1 Empirical price pressure and inventory dynamics

Markets where investors do not continuously monitor and participate deviate from the standard Walrasian tradition (Townsend (1978), Grossman and Miller (1988), Rust and Hall (2003), and many others). Historically, the NYSE assigns one intermediary, called a specialist, to act as an market maker for each stock. This structure clearly identifies the amount of liquidity supplied by an intermediary. We use data on these intermediaries' inventory positions to identify price pressure both in the cross-section and through time. While the designation of a single intermediary is relatively unique to the NYSE, the fundamental economic forces that generate prices pressure and intermediary inventory risk exist in all markets where investor trading needs are not perfectly synchronized. Section 3 further discusses the NYSE structure and changes to that structure subsequent to our sample period. Before any estimations we discuss the data and provide some summary statistics.

1.1 Data and summary statistics

We use data from the Center for Research in Security Prices (CRSP), the NYSE's Trade and Quotes (TAQ), and a proprietary NYSE dataset to construct the end-of-day midquote price (i.e., the average of the bid price and ask price) and NYSE specialist inventory position along with other variables from 1994 through 2005. We construct a balanced panel to make results comparable through time and control for stock fixed effects. We start with the sample of all NYSE common stocks that can be matched across TAQ and CRSP and retain the stocks that are present throughout the whole sample period. Stocks with an average share price of less than \$5 are removed from the sample, as are stocks with an average share price of more than \$1,000. The resulting sample comprises 697 common stocks. Stocks are sorted into quintiles based on market capitalization. Quintile 1 refers to the large-cap stocks and quintile 5 corresponds to the small-cap stocks. To facilitate comparisons across stocks we convert the NYSE specialist position, which is in shares in the original database, into dollars by multiplying the position times each stock's average price each year.[6] This eliminates daily price changes in the inventory variable allowing its use as an explanatory variable for the transitory price effect in the econometric model. Throughout we use specialist, market maker, and intermediary interchangeably to refer to the NYSE specialist.

[insert Table 1]

Summary statistics. Table 1 presents the mean of various trading variables by size quintile. To provide a sense of the variable's variability through time the within variation, which is defined as the standard deviation of the data series after removing stock fixed effects, is also included. Several observations emerge from the statistics. First, the within standard deviation in intermediary inventory is \$1.4 million, which is substantial relative to her average position. It suggests that the specialist is an active intermediary in matching buyers and sellers at interday horizons. Table 2 subsequently disaggregates this variation by year and by size quintile. Second, while not the focus of this study, the average position of the intermediary is positive and economically significant.[7] For example, for the large-cap stocks in Q1 she maintains an almost half a million dollar average inventory position. The inventory position for the small-cap stocks in Q5 is considerably smaller at \$77,900. Third, the market capitalization is \$34 billion for Q1 stocks and declines to \$290 million for Q5 stocks. The effective spread (more precisely referred to the half spread, but for notational convenience we use spread throughout), which is defined as the distance between the transaction price and the prevailing midquote price, is 8 basis points for Q1 stocks and 46 basis points for Q5 stocks, demonstrating considerable heterogeneity across stocks.

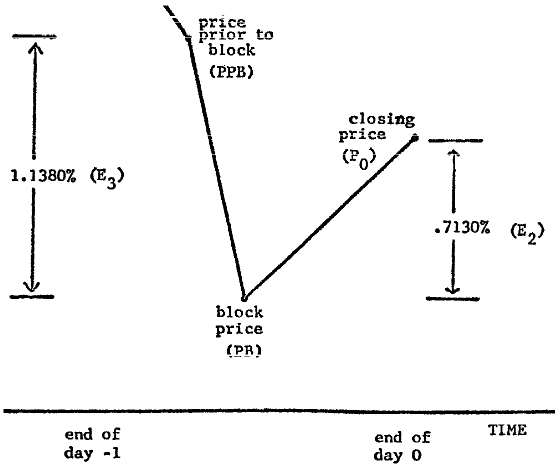
1.2 State space model to disentangle price pressures and efficient price innovations

One challenge in identifying price pressures is that investors' net order imbalance, the difference between the volume investors buy and sell (which equals the intermediary's inventory change), may convey information as well as cause pressure which makes prices 'overshoot'. This well-known pattern has been documented in various event studies. For example, Kraus and Stoll (1972) show

[6] We use the stock-split and dividend information from CRSP to remove these effects from the midquote prices in TAQ.

[7] This average long position is likely driven by a combination of a cost asymmetry between long and short position and capital requirements imposed by the exchange which intermediaries choose invest in stocks. The model in Section 2 interprets the zero position as the deviation from the long-term optimal position of the intermediary.

that prices overshoot in the event of a block trade, i.e., a transaction where the initiating party wants to trade an significant number of shares, typically defined as greater than 10,000. They document the following pattern for block sales where E_3 is the immediate total impact of the block sale, $E_3 - E_2$ is the permanent impact, and E_2 is the temporary impact (price pressure) from the sale:



source: Kraus and Stoll (1972, p.575)

To identify the price pressure effect in the presence of an information effect, we use the state space approach of Menkveld, Koopman, and Lucas (2007) which models an observed, high-frequency price series as the sum of two unobserved series: a nonstationary efficient price series ('information') and a stationary series that captures transitory price effects ('price pressures'). We use log prices throughout the paper and remove a required return by subtracting a linear trend with a slope equal to the riskfree rate plus beta times a market risk premium of 6%. In its simplest form the model structure for the detrended log price is:

$$p_t = m_t + s_t, \quad (1)$$

$$m_t = m_{t-1} + w_t, \quad (2)$$

where p_t is the observed price, m_t is the unobserved efficient price, s_t is the unobserved transitory price effect, and w_t is the innovation in the efficient price. s_t and w_t are mutually uncorrelated and normally distributed. It is immediate from the structure of the model that only draws on w_t affect the security's price permanently as any draw on s_t is temporary as it affects prices only at a single point in time. The model can be estimated with maximum likelihood where the likelihood is calculated by the Kalman filter. Details on the implementation are in Appendix I.

There are several reasons why the state space methodology is preferable to other approaches such as generalized method of moments (GMM) or an autoregressive moving average (ARIMA) model. First, maximum likelihood estimation is asymptotically unbiased and efficient. Second, the model implies that the differenced series is an invertible first-order moving average (MA(1)) time series model which implies a infinite lag autoregressive (AR) model. This is particularly cumbersome if the price series (or the inventory series that we will use as explanatory variable for s_t) have missing values. The Kalman filter ensures maximum efficiency in dealing with missing values by not losing any information as it considers the likelihood of *all* level series changes even if they extend over multiple periods in the case of missing observations. Any method based on the differenced series does not consider that information. Third, after estimation, the Kalman smoother (essentially a backward recursion after a forward recursion with the Kalman filter) facilitates a series decomposition where at any point in time the efficient price and the transitory deviation are estimated using past and future prices. This allows for an in-sample decomposition of prices which we will illustrate. Durbin and Koopman (2001) provide an extensive treatment on the use of state space models to analyze economic times series.

The remainder of this subsection develops the general state space model to be taken to the data. We first develop the latent efficient price process and then the stationary price deviations that should capture the price pressures that are our focus.

Unobserved efficient price process. We use the model to analyze daily midquote price series by stock-year. The efficient price series is a martingale that consists of two components:

$$m_{it} = m_{i,t-1} + \beta_i \hat{f}_t + w_{it}, \quad w_{it} = \kappa_i \hat{I}_{it} + u_{it} \tag{3}$$

where the subscript i indexes stocks, t indexes days, $\hat{f}_t$ is a common factor innovation that is obtained as the residual of an autoregressive time series model applied to the market cap weighted average of standardized midquote returns (we standardize to control for heteroscedasticity), w_{it} is the idiosyncratic innovation, $\hat{I}_{it}$ is the idiosyncratic inventory innovation (similar procedures are used as described for $\hat{f}_t$) that represents the 'surprise' net order imbalance which is potentially informative, and u_{it} is the stock-specific innovation orthogonal to this order imbalance innovation and assumed to be a normally distributed white noise process.[8] The decomposition of efficient price innovation into a common factor component ($\beta_i \hat{f}_t$) and an idiosyncratic component (w_{it}) is relevant for our purposes as the latter represents undiversifiable risk for the intermediary. The common factor risk is easily hedged through highly liquid index products. For the same reason we remove the common factor from inventory dynamics as the price risk over a market-wide shock to inventory is similarly easily hedged through index products.

Unobserved transitory price deviations process. We propose the following process for the stationary price deviations:

$$s_{it} = \alpha_i I_{it} + \beta_i^0 \hat{f}_t + \cdots + \beta_i^k \hat{f}_{t-k} + \varepsilon_{it} \tag{4}$$

where the error term ε_{it} is normally distributed and uncorrelated with w_{it}. Inventory enters as an explanatory factor to allow price pressure to originate from the intermediary's desire to mean revert inventory (Section 2 provides a simple dynamic model generating such a linear structure). The f terms enter to capture a documented lagged adjustment to common factor innovation that is particularly prevalent for less actively traded small-stocks (see, e.g., Campbell, Lo, and MacKinlay (1997)). In the proposed specification the beta coefficient in the efficient price process captures the long-term impact of a common factor shock on the price of the security and any lagged adjustment shows up through negative beta coefficients in the transitory price effects equation.

Observed price process. We close the econometric model with the observation equation:

$$p_{it} = m_{it} + s_{it}. \tag{5}$$

As we intend to analyze price pressures in the cross-section as well as in the time dimension all empirical analysis are done by stock-year. To report the 697*12 stock-year results we aggregate stocks into quintile bins according to their size. The allocation across size quintiles is fixed throughout the sample so as to ensure that bins are comparable across years. Means are calculated for each bin along with the number of t-statistics that are outside of the 0.10 to 0.90 quantile interval. These t-statistics are available as supplementary material on the authors' websites. The

[8]While not the focus of the paper, it is worth noting that the idiosyncratic inventory innovation term in the martingale equation eliminates a potential 'omitted variable bias' as the explanatory variable for the efficient price innovation (w_t) correlates with the explanatory variable for price pressure (s_t), i.e., the order imbalance innovation (unexpected change in inventory) in period t correlates with inventory at time t (which is at the end of period t).

tables report the p-value of a meta test statistic that counts the number of significant t-values in a size-year bin (and in the aggregation across bins). This statistic is binomially distributed where the probability of 'success' equals the significance level of the t-test performed for each stock-year estimation, i.e., 0.20 in our two-tailed test.[9]

Time series statistics as preliminary evidence on price pressures. Before turning to the state space model estimates, some straightforward time series statistics are useful to test whether the signs of the price effects that we are after are present in the data.

[insert Table 2]

Panels A and B of Table 2 reports first and second order autocorrelation of idiosyncratic midquote returns. The effects of contemporaneous and lagged adjustment to the common factor innovation are removed by regressing the midquote return on the common factor innovation up to four lags. The residuals serve as the idiosyncratic returns. Consistent with the individual stocks autocorrelation results in Campbell, Lo, and MacKinlay (1997) the average first order autocorrelation is negative in 64 of the 70 size-year bins. The low p-values indicate that the coefficient estimates are significant at conventional significance levels.[10] The negative first order autocorrelation is consistent with transitory price effects as the simple state space model of equations (1) and (2) implies a negative first-order correlation in midquote returns. The table further shows that the second order autocorrelation is also significantly negative which indicates that price pressures may carry over days rather than being an ultra-high-frequency intraday phenomenon. This also implies that the unconditional transitory deviations are potentially much larger relative to fundamental volatility than the simple first-order autocorrelations suggest.[11]

Panels C, D, and E of Table 2 further reports the standard deviation of intermediary inventory, its autocorrelation, and a cross-correlation with subsequent midquote returns as preliminary evidence on the conjectured relation between transitory price deviations and inventory. The standard

[9] Correlation across stocks is limited by the empirical analysis' focus on idiosyncratic effects by removing a common factor in both the price and in the inventory series.

[10] Overall, 3525 t-values are significant and 4836 t-values are insignificant. The sign of the significant t-values is primarily negative (2427 negative vs. 1098 positive) which shows that the negative means are statistically significant. We omit extensive statistical significance discussion in the remainder of the document for brevity.

[11] If the stationary term follows an AR(1) process, $s_t = \varphi s_{t-1} + \varepsilon_t$, the first-order autocorrelation in midquote returns is:

$$\rho_1 = \frac{-(1-\varphi)\sigma_\varepsilon^2}{(1+\varphi)\sigma_w^2 + 2\sigma_\varepsilon^2}$$

where φ is the autoregressive coefficient in the s_t process. The more persistent the price pressure, higher φ, the less negative the return autocorrelation becomes.

deviation of intermediary end-of-day inventory is \$1.131 million for the large-cap stocks and monotonically decreases to \$165,000 for the small-cap stocks. It is relatively constant throughout time but tapers off in the last few years in the sample. The cross-sectional variation is undoubtedly due to a higher fundamental volatility and a smaller less active market for small-cap stocks. In this case, the intermediary will shy away from frequent and large nonzero inventory positions, e.g., see the model proposed in Section 2.

The inventory volatility is roughly twice the mean inventory across all quintiles. This is evidence of active intermediation across days as inventory management is not a phenomenon that is restricted to the intraday ultra-high frequency level where intermediaries 'go home flat.' The table further documents a significant first order autocorrelation in inventories which show that these positions can last for multiple days. Again, there is considerable cross-sectional variation as the average autocorrelation for the large-cap stocks is 0.28 which monotonically increases to an average autocorrelation of 0.72 for small-cap stocks, corresponding to inventory half-lives of 0.55 days for the largest stocks and 2.11 days for the smallest stocks.[12] The intermediary seems to trade out of most of an end-of-day position in the course of the next day for the large-cap stocks, whereas it take multiple days for the small-cap stocks. Finally, we calculate the correlation between today's inventory position and tomorrow's midquote return to verify whether the two sets of results in the table can be reconciled. The last panel in the table shows that today's inventory position correlates significantly with tomorrow's midquote return. The positive signs are consistent with models of intermediary risk management as the intermediary applies price pressure by lowering the midquote price on a long position (relative to the long-term average) which elicits an order imbalance (more investor buying than selling) that mean reverts her inventory and, as a consequence, leads to smaller price pressure. This creates a positive correlation between today's position and tomorrow's midquote return.

State space model estimates. Before presenting the general state space model defined by equations (3), (4), and (5), we illustrate the idea of the model graphically. The equations relate the permanent price change Δm_t and transitory price effect s_t to the intermediary's inventory series.

[12] These inventory autocorrelations are lower than the puzzlingly large autocorrelations found in NYSE data from the late 1980s and early 1990s in Madhavan and Smidt (1993) and Hasbrouck and Sofianos (1993).

Initially, we do not use inventory data to estimate the reduced version of the model:

$$p_{it} = m_{it} + s_{it} \tag{6}$$

$$m_{it} = m_{i,t-1} + \beta_i \hat{f}_t + w_{it} \tag{7}$$

$$s_{it} = \varphi_i s_{i,t-1} + \beta_i^0 \hat{f}_t + \cdots + \beta_i^3 \hat{f}_{t-3} + \varepsilon_{it} \tag{8}$$

where the AR(1) process for s_{it} allows transitory price effects to be persistent as suggested by the negative second order auto correlations in returns (Panel B of Table 2) and the inventory persistence (Panel D of Table 2). Model estimates by stock-year are in the supplementary material.

[insert Figure 1]

Figure 1 illustrates the model estimates for twenty trading days in a representative stock (Rex Stores Corporation, ticker RSC, CRSP PERMNO 68830) starting January 14, 2002. It exploits one attractive feature of the state space approach, which is that conditional on the model's parameter estimates the Kalman smoother generates estimates of the unobserved efficient price and the transitory price pressure processes conditional on *all* observations. In other words, it uses past and future prices to estimate the efficient price m_{it} and the temporary price deviation s_{it} at any time t in the sample. The first graph plots the observed midquote price and the efficient price estimate. It illustrates that price deviations from fundamental value are economically large—hundreds of basis points—and persistent as they appear to last for multiple days. The second graph plots the differential between the observed price and the efficient price estimate, i.e., the price pressure, along with the intermediary's inventory deviation from its long-term mean. The clear negative correlation between the two series is consistent with the intermediary using price pressure to mean-revert her inventory towards its desired level. The third graph plots the innovation in efficient price Δm_t against the contemporaneous 'surprise' idiosyncratic inventory change which is obtained as the residual of an AR(9) model. It indicates that unexpected order flow is informative on the efficient price. A surprise positive inventory change indicates unexpected selling by liquidity demanders which changes the efficient price downwards as the selling might have been driven by information. These observations are now tested rigorously by estimating the state space model with inventory data.

[insert Table 3]

Table 3 reports the estimates of the general state space model defined by equations (3), (4), and (5). In Panel A the key parameter α_i that measures the conditional price pressure has the conjectured negative sign and is statistically significant as 4045 t-values are significantly negative, 4051 are insignificant, and only 265 are significantly positive. Prices are low when the intermediary is on a long position and high when she is on a short position relative to her long-term mean inventory. Conditional price pressure exhibits substantial cross-sectional variation as α_i is -0.02 for the large-cap stocks and -1.01 for the small-cap stocks. These numbers are economically significant as a \$1.131 million (one standard deviation) position change in intermediary inventory creates a price pressure of 1131*0.02=17 basis points (cf. effective half-spread of 8 basis points, see Table 1). A similar position change in the small-cap stocks would create a price pressure of 1131*1.01=1142 basis points or 11.42%!

Panel B of Table 3 further shows that the average price pressure varies less in the cross-section than the conditional price pressure. The average pressure which is measured as the conditional pressure times the standard deviation of inventory is 17 basis points for the large-cap stocks and 120 basis points for the low-cap stocks. The conditional pressure is roughly 7 times higher for the small-cap stocks relative to the large-cap stocks whereas the conditional pressure is 50 times higher with the difference attributable to the intermediary taking smaller positions in the smaller stocks. Panel F also reports the size of these average price pressures relative to permanent volatility. The variance ratio is 0.02 for the large-cap stocks which indicates that these price pressures are small relative to fundamental volatility. In contrast the small-cap stock variance ratio is 1.32, implying transitory volatility due to price pressures is larger than fundamental volatility. Finally, the average price pressure variance identified using inventories is almost half of the transitory price deviations variance in the model that does not use inventory data (equations (6)-(8), see supplementary material for estimation details). The ratio is $(49/75)^2 = 0.42$ which could be interpreted as an 'R^2'.

Consistent with earlier empirical work (for example, Hasbrouck (1991)), unexpected inventory changes, which equal unexpected investor buying or selling, explain a significant part of efficient price innovations. The κ_i estimates in Panel C are -0.35 on average and highly significant as 6602 t-values are significantly negative, 1631 are insignificant, and only 128 are significantly positive. In Panel D the average explained permanent volatility—κ_i times the standard deviation of unexpected inventory changes—is 56 basis points. This compares to an average total permanent volatility of

177 basis points in Panel E.

2 A simple dynamic inventory control model to interpret the time series properties of price pressure

This section models the intermediary's dynamic inventory control policy as a 'stochastic optimal linear regulator' (SOLR) problem (see Ljungqvist and Sargent (2004, p.112)). A representative intermediary faces stochastically arriving investors with elastic liquidity demands. The solution characterizes a stationary distribution for price pressures which along with the empirical results in the previous section, enables identification of the intermediary's relative risk aversion and liquidity demanders' private value distribution. It also allows for an analysis of the costs of trading and social welfare. Given that the model assumes investors arrive asynchronously, first-best is not achievable, so our social welfare analysis focuses on deviations from constrained Pareto efficiency; for expositional ease we will often omit constrained when discussing Pareto efficiency.

To fit the SOLR approach we assume that: (i) liquidity demand which determines the intermediary's buy and sell volume is exogenous and normally distributed with a mean that is linear in the bid and ask price, respectively, (ii) the intermediary is a mean-variance optimizer over nonstorable consumption,[13] and (iii) a security position exposes the intermediary to fundamental value risk which is modeled as a normally distributed stochastic dividend to avoid a notational burden.[14] The fundamental value of the security, m_t in the empirical model, is set to zero.

The model is inspired by Ho and Stoll (1981) who frame dynamic inventory control in a standard macro model of a CRRA utility intermediary who controls the public buy and sell rate which are linear in her ask and bid price quotes. The intermediary solves a dynamic program to maximize terminal wealth. We deviate by setting up the problem as an infinite horizon one so as to generate a stationary distribution for price pressure that can be compared to the empirical estimates.

[13] Cochrane (2001, p.155) also considers quadratic utility a natural starting point when he introduces dynamic programming. Madhavan and Smidt (1993) model inventory as having quadratic holding costs. Lagos, Rocheteau, and Weill (2009) model liquidity provision where intermediaries maximize over nonstorable consumption.

[14] The stochastic dividend minimizes accounting in the model, but, as Ho and Stoll (1981, p.52) emphasize, no sources of uncertainty are ignored. These assumptions have several implications. First, the nonstorable consumption removes the ability of the intermediary to precautionary save and smooth consumption. Second, quadratic utility leads to risk aversion that increases with wealth. Third, the normality assumption for public buy and sell volume implies that they could become negative. To show that the model's predicted price pressure dynamics is robust, we also numerically solve an infinite horizon 'Ho and Stoll' model in the supplementary material.

2.1 The environment

Time is discrete and infinite. There is one durable asset that produces a perishable and stochastic (numéraire) consumption good. An infinitely-lived intermediary supplies liquidity by standing ready to buy or sell the asset to outside investors who demand liquidity. At the start of each period a single intermediary quotes an ask price at which she commits to sell and a bid price at which she commits to buy.

Reduced form liquidity demand. Liquidity demand is characterized by the primitive parameters λ, which is the aggregate amount of all investors' private values to trade per unit of time, and θ, which captures the dispersion of private values across investors. Investor buying and selling are determined by the intermediary's bid price, $s - \delta$, and ask price, $s + \delta$, which are characterized by price pressure s, which is $p_t - m_t$ in the state space model, and the bid-ask spread of 2δ. Specifically, investor buying and selling demand for liquidity are normally distributed variables linear in the bid-ask quotes:[15]

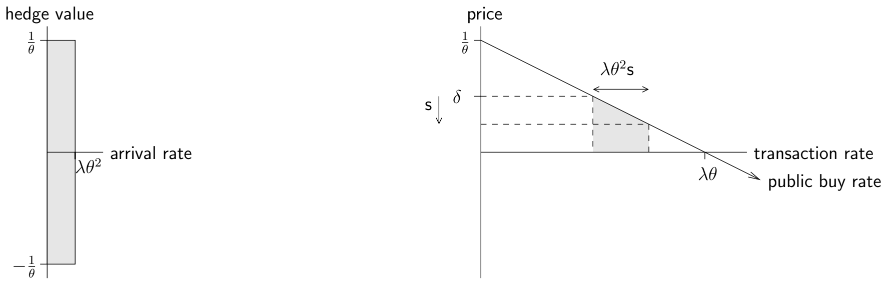
$$q_s(s,\delta) = \lambda\theta\Big(1 - \theta(s+\delta)\Big) + \varepsilon_s, \quad \varepsilon_s \sim \mathrm{N}\Big(0, \frac{1}{2}\sigma_\varepsilon^2\Big), \tag{9}$$

$$q_b(s,\delta) = \lambda\theta\Big(1 + \theta(s-\delta)\Big) + \varepsilon_b, \quad \varepsilon_b \sim \mathrm{N}\Big(0, \frac{1}{2}\sigma_\varepsilon^2\Big), \tag{10}$$

where ε_s and ε_b are independent of each other and identically and independently distributed (IID) each period. The environment is best illustrated by the two graphs below where graph (i) illustrates liquidity demand 'distribution' over private values and graph (ii) depicts how such demand translates into an expected investor buy rate, which equals the intermediary sell rate, that is linear in the ask price.

The intermediary's liquidity supply. The intermediary is risk-averse and therefore dislikes the risky dividend associated with a nonzero inventory position. She solves the following infinite-

[15] These distributions can arise from an economy where each time period each of a finite number N of investors each owning $\frac{1}{N}$ units of the asset receives a shock to their private value to trade with probability $2\lambda\theta$. Common interpretations of such shocks are that investors have hedging or liquidity needs from which they derive private values (Grossman and Miller (1988)). Private value shocks are uniformly distributed on support $[-\frac{1}{\theta}, \frac{1}{\theta}]$. If the intermediary quotes an ask price of $(s+\delta)$ investors' sell volume is distributed $\frac{1}{N}$ times a binomial random variable with parameters $(N, \lambda\theta(1-\theta(s+\delta)))$. The binomial distribution for intermediary sell volume can then be approximated by a normal distribution with the same mean but with variance equal to $\frac{1}{2}\sigma_\varepsilon^2$. Buy volume is similarly distributed.



(i) mean demand per private value (ii) mean public buy volume as function of ask price

horizon dynamic program:

$$v_{i_0} = \max_{\{s_t(i^t),\delta_t(i^t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \mathrm{E}_{i_0}\Big(\beta^t((c_t - \frac{1}{2}\lambda) - \frac{1}{2}\tilde{\gamma}(c_t - \frac{1}{2}\lambda)^2)\Big) \tag{11}$$

where i^t represents the history of her inventory position through time t and $\mathrm{E}_{i_0}(.)$ is the expectation operator conditional on starting off with inventory position i_0. The quadratic utility parametrization is such that the Arrow-Pratt coefficient of (absolute) risk aversion (ARA) is $\tilde{\gamma}$ at $\frac{1}{2}\lambda$ which is the expected consumption of a monopolistic risk-neutral intermediary. The intermediary's actual consumption in period t equals net trading revenue plus the stochastic dividend:

$$\begin{aligned} c_t &= (s_t + \delta_t)q_s(s_t, \delta_t) - (s_t - \delta_t)q_b(s_t, \delta_t) + i_t \Delta m_{t+1}, \qquad (12) \\ &= 2\lambda\theta(\delta_t - \theta(s_t^2 + \delta_t^2)) + s(\varepsilon_{st} - \varepsilon_{bt}) + \delta(\varepsilon_{st} + \varepsilon_{bt}) + i_t \Delta m_{t+1}, \quad \Delta m_{t+1} \sim \mathrm{N}(0, \sigma^2) \end{aligned}$$

where Δm_{t+1} is the stochastic dividend at the start of period $t+1$ (which runs from t to $t+1$). Consumption mean and variance therefore equal:

$$\mathrm{E}(c_t) = 2\lambda\theta(\delta_t - \theta(s_t^2 + \delta_t^2)) \tag{13}$$

$$\mathrm{var}(c_t) = \sigma_\varepsilon^2(s_t^2 + \delta_t^2) + \sigma^2 i_t^2 \tag{14}$$

The law of motion for inventory follows from the net trade in the asset:

$$i_{t+1} = i_t - q_s(s_t, \delta_t) + q_b(s_t, \delta_t) = i_t + 2\lambda\theta^2 s_t - \varepsilon_{st} + \varepsilon_{bt} \tag{15}$$

A final step simplifies the problem in two ways. First, the expected utility expression is linearized in mean and variance by a first-order Taylor expansion around the average risk-neutral consumption $\frac{1}{2}\lambda$. This removes the quadratic conditional mean which would make the objective function fourth order and therefore impossible to solve with standard techniques. Second, we omit the impact of σ_ε^2 in the variance of consumption, equation (14), to focus on how prices reflect the dynamic trade-off between the intermediary's expected cost of a pressured price and its benefit of mean-reverting risky inventory. Omission of the consumption variance effect is innocuous if the pain of liquidity demand uncertainty (as captured by $\frac{1}{2}\tilde{\gamma}\sigma_\varepsilon^2$) is small relative to the expected revenue loss due to price pressure (as captured by $2\lambda\theta^2$). The final specification of the dynamic program is therefore:

$$v_{i_0} = \max_{\{s_t(i^t),\delta_t(i^t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \mathrm{E}_{i_0}\left(2\lambda\theta\delta_t - 2\lambda\theta^2(s_t^2+\delta_t^2)) - \frac{1}{2}\tilde{\gamma}\sigma^2 i_t^2\right) \tag{16}$$

$$i_{t+1} = i_t + 2\lambda\theta^2 s_t + \varepsilon_t, \quad \varepsilon_t \sim \mathrm{N}(0,\sigma_\varepsilon^2). \tag{17}$$

2.2 The recursive form, the closed-form solution, and its characteristics

The IID character of net liquidity demand uncertainty (ε_t) and stochastic dividend (Δm_{t+1}) allows for a recursive formulation of the dynamic program. The intermediary solves the following Bellman equation:

$$v_i = \max_{\{p,\delta\}} \mathrm{E}_{i_0}\left(2\lambda\theta\delta - 2\lambda\theta^2(s^2+\delta^2) - \frac{1}{2}\tilde{\gamma}\sigma^2 i^2 + \beta v_{i'}\right) \tag{18}$$

$$i' = i + 2\lambda\theta^2 s + \varepsilon. \tag{19}$$

Standard solution techniques yield the following solution:

$$v_i = \frac{\lambda}{2(1-\beta)} - P(i^2 + \frac{\beta}{1-\beta}\sigma_\varepsilon^2) \tag{20}$$

$$s^* = \alpha i, \quad \alpha \equiv \frac{-1}{\frac{1}{\beta P}+Q} \tag{21}$$

$$\delta^* = \frac{1}{2\theta} \tag{22}$$

where (s^*, δ^*) denote the optimal prices controls and

$$P \equiv \frac{-(1-\beta)+\beta RQ + \sqrt{(1-\beta)^2 + 2\beta(1+\beta)QR + \beta^2Q^2R^2}}{2\beta Q}, \quad Q \equiv 2\lambda\theta^2, \quad R \equiv \frac{1}{2}\tilde{\gamma}\sigma^2. \tag{23}$$

Time series properties of price pressure. The solution to the intermediary's control problem implies linear dynamics for price pressure (s) and inventory which matches the econometric model that was taken to the data. More specifically, price pressure is linear in inventory (equation (21)) and inventory is characterized by first-order autoregressive process:

$$i_t = \frac{\beta PQ}{1 + \beta PQ} i_{t-1} + \varepsilon_t = -Q\alpha i_{t-1} + \varepsilon_t. \tag{24}$$

Orthogonality of the spread and optimal inventory control. The intermediary's pricing strategy decomposes into two orthogonal strategies. First, she sets a spread that exploits her monopolistic pricing power vis-à-vis the investors. Second, she uses price pressure (the midquote price) to optimally mean-revert her inventory. Mathematically, the control strategy orthogonalizes as the spread does not enter the speed of mean-reversion and the spread is constant across inventory states.[16] This critically depends on the assumption that public buy and sell volume are linear in prices (see equation (9) and (10)).

The orthogonality result implies (i) that price pressure, not spread, should be the focus of a study on dynamic inventory control by an agent who intermediates nonsynchronous buy and sell volume and (ii) that a competitive equilibrium can be characterized. If the 'representative' intermediary becomes a price-taker and entry drives rents to zero, then the candidate pricing function is equal to the price pressure exercised by the monopolistic intermediary with a spread exactly compensating for the cost of the stochastic dividend on nonzero inventory positions. From this we can calculate the model-implied competitive spread and compare it with the observed spread.

2.3 Identification of the model's primitive parameters

The remainder of the section uses the state space model estimates to identify the model's primitive parameters, calculates an model-implied competitive spread and finally, in the next subsection, analyzes Pareto efficiency.

[insert Table 4]

Table 4 summarizes all results with a repeat of the empirical results in Table 3 needed for identification of the model's primitive parameters. To separately identify the λ and θ parameters,

[16] Zabel (1981) and Mildenstein and Schleef (1983) find the spread is independent of inventory. Ho and Stoll (1981) find inventory has a very small effect on the spread.

which characterize investors' demand for liquidity, trading volume and spread are needed. See Appendix II for detailed calculations. The only additional parameter needed is discount factor (β) which is assumed to be equal to the risk-free rate.

Competitive spread required to compensate for inventory risk. Equation (20) shows that the value of being a liquidity supplier naturally decomposes into a discounted expected revenue due to earning the bid-ask spread (first term) and a discounted cost due to price risk associated with nonzero inventory (second term). The annuity value of the inventory risk compensation equals $\beta P \sigma_\varepsilon^2$ if the intermediary starts with zero inventory (which is assumed for the remainder of the section). The model-implied conditional price pressure and inventory mean-reversion as described by equations (21) and (24) identify the factor βPQ. The remaining factor σ_ε^2 follows directly from inventory variance and mean-reversion, i.e., $\sigma_\varepsilon^2 = (1 - \rho_I)\sigma^2(I)$, where ρ_I is the first-order autocorrelation of inventory. This daily value divided by the intermediated daily volume naturally defines the 'competitive spread'.

Panel B of Table 4 reveals that the median model-implied competitive spread is 7.32 basis points for the largest-cap stocks and monotonically increases to 73.16 basis points for the smallest-cap stocks.[17] These results do not use actually observed spread, but are identified solely from conditional price pressure and inventory dynamics. It is reassuring that the model-implied spread is of the same magnitude as the observed spread. It is slightly below for largest-cap stocks (7.32 vs. 8.41 basis points) and increases to a roughly 50% higher for the smallest-cap stocks (73.16 vs. 46.12 basis points). The differential is potentially due to a specialist privilege during the sample period where she could see the incoming order flow and limit her trading with privately-informed order flow (see, e.g., Rock (1990)). This privilege is particularly valuable for small-caps which are relatively more opaque. The observed spread is net of the value of such privilege (given that access to specialist seats is competitive) and, therefore, can be lower than the competitive spread.

The model's primitive parameters. All primitive parameters are identified if in addition to the terms identified in the previous section (βP, Q, and σ_ε^2), one uses the reciprocal of the gross riskfree rate (from Kenneth French's website) for the discount factor β, uses the empirical random-walk volatility ($\sigma(w)$) for the fundamental price risk (σ), and uses the observed empirical

[17] For the remainder of the section medians and interquartile ranges ($Q_{0.75} - Q_{0.25}$) are reported as the model-implied variables are nonlinear functions of the empirical estimates (with estimates in denominators) which creates numerical problems. Medians and interquartile ranges are robust to these issues.

results on intermediated daily volume ($specialist\_particip * daily\_volume$) and the effective spread ($espread$).[18] Appendix II summarizes all variables identified in this section and expresses them in terms of the previous section's empirical estimates. The private value rate λ is highest for the largest-cap stocks: \$56,690 per stock per day. It decreases monotonically to \$1,460 for the smallest-cap stocks. Private value dispersion $\frac{1}{\theta}$ is 171 basis points for largest-cap stocks, 135 for second largest-cap stocks and then increase monotonically to 307 basis points for smallest-cap stocks.

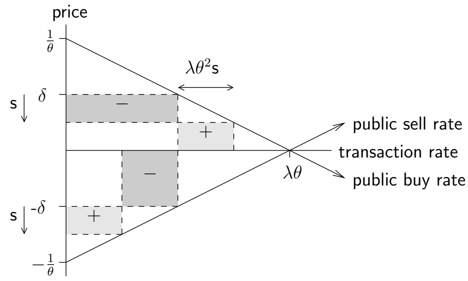
We calculate the coefficient of relative risk aversion (γ) by dividing $\tilde{\gamma}$ by the intermediaries expected consumption at the competitive spread. The median relative risk aversion declines monotonically from 0.24 for the largest-cap stocks to 0.05 for the smallest-cap stocks with the differences across quintiles small relative to the interquartile ranges within each quintile. The overall median of 0.10 is low relative to relative risk aversion estimated from asset, insurance, and labor markets (e.g., Mehra and Prescott (1985), Barsky, Juster, Kimball, and Shapiro (1997), Chetty (2006), and Cohen and Einav (2007)). The low risk aversion could arise from risk tolerant capital migrating to the intermediation sector or be a sign of an agency conflict between firm and the individual traders it employs to act as intermediaries. If the traders have limited liability, then they may be willing to take more risky bets than the firm's owners would prefer. It is also possible that our structural model underestimates true risk aversion as the model ignores increased net demand uncertainty associated with price pressure (see related discussion in Section 2.1).

2.4 Deadweight loss due to Pareto inefficiency

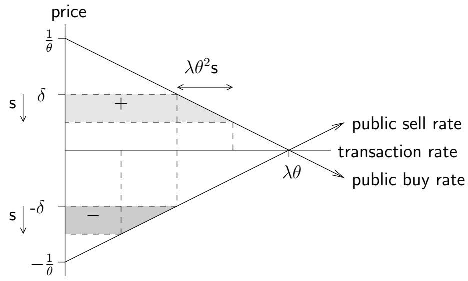
The intermediary trades off the expected cost of price pressure against the benefit of inventory mean-reversion. The left-hand side graph in the figure below depicts the expected cost of a negative price pressure s (due to a positive inventory). The ask price is lowered and the light gray area above the horizontal axis indicates the additional expected revenue due to more public buying. The lowered price also reduces the margin on all buys which is indicated by an expected loss equal to the dark gray area above the axis. Similar areas are drawn for the lowered bid price. Overall, the differential between the light and dark gray areas is the expected revenue decline due to price pressure. The size of this area equals the $2\lambda\theta^2 s^2$ cost in the intermediary's Bellman equation (see equation (18)).

[18] The identification of λ and θ relies primarily on trading volume and is relatively insensitive to the spread.

The right-hand side graph shows that the liquidity demanders benefit as the discount they enjoy when buying exceeds the lower price they receive when selling. The difference equals half the intermediary's cost of price pressure: $\lambda\theta^2 s^2$. The daily annuity value of the net subsidy is obtained recursively (see Appendix III) and, when divided by daily intermediated volume, equals 0.58 basis points for the largest-cap stocks (8% of their competitive spread) and 15.52 basis points for the smallest-cap stocks (21% of their competitive spread).

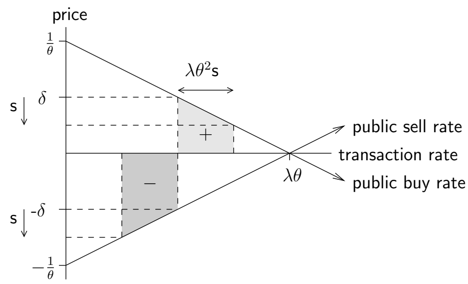


(i) intermediary's cost of price pressure



(ii) liq. demander benefit of price pressure

The intermediary only experiences the cost of lost revenue due to price pressure and does not internalize the benefit it creates for liquidity demanders which, most likely, makes allocations Pareto inefficient.[19] The adjective 'constrained' is used here to emphasize that the first-best of synchronous arrivals eliminating the need for intermediation is not attained. A social planner might Pareto improve by making the intermediary suffer the *net* surplus destroyed by price pressure rather than her private revenue loss (i.e., the sum of graphs (i) and (ii) rather than graph (i)). This true social cost is $\lambda\theta^2 s^2$ as depicted in the following graph. Solving the 'social planner dynamic program'



social cost of price pressure

[19]See Weill (2007) and Lagos, Rocheteau, and Weill (2009) for models with conditions where the intermediation sector does and does not achieve Pareto efficient.

yields a solution similar to the intermediary's solution where P in equation (23) is replaced by $\tilde{P}$:

$$\tilde{P} = \frac{-(1-\beta) + 2\beta RQ + \sqrt{(1-\beta)^2 + 4\beta(1+\beta)QR + 4\beta^2 Q^2 R^2}}{4\beta Q} \tag{25}$$

Panel C of Table 4 shows that the social planner's implied 'competitive spread' is 6.45 basis points for the largest-cap stocks and increases to 54.77 basis points for the smallest-cap stocks. The gain relative to the net spread paid by the liquidity demander in the intermediary's solution (i.e., net of the subsidy) is 0.05 basis points (0.8% of the competitive spread) for Q1 stocks and 2.58 basis points (4.5% of the competitive spread) for Q5 stocks.

3 Price pressure and NYSE market structure

When interpreting our results it is worth discussing the institutional structure of the specialist intermediary at the NYSE. Historically, the NYSE granted the specialists a central position in the trading process and imposed obligations upon the specialists. Panayides (2007) shows that the most significant obligation, the Price Continuity Rule which 'requires the specialist to smooth transaction prices by providing extra liquidity as necessary to keep transaction price changes small,' is important at the transaction-level horizon. Panayides finds that the rule causes specialists to accumulate inventory to prevent 'transaction prices from overshooting beyond their equilibrium levels'. This causes inventories to be positively associated with transitory price effects, the opposite of our relation between intermediary inventory and price pressure. Therefore, if the Price Continuity Rule manifests itself at a daily frequency it causes underestimation of price pressures associated with inventory.

While the NYSE designates a single intermediary for each stock, it is possible for other investors to compete with the specialist by placing limit orders to supply liquidity. Such a possibility is especially important given the NYSE's recent market structure changes (after our sample period) which resulted in a reduced role for the specialist (Hendershott and Moulton (2007)). Additional liquidity suppliers could reduce the width of the bid-ask spread and could also reduce the social costs of intermediaries bearing risk. The most efficient manner to share risk is for the inventory to be immediately and equally shared across all liquidity suppliers, leading to perfectly correlated positions. How would this affect our estimates of price pressure? First, the conditional price pressure per unit of inventory (α) should be adjusted by the specialist's fraction of inventory, e.g., if the specialist carries one half of the total inventory then the conditional price pressure should be multiplied by one half. The average price pressure ($\alpha\sigma(I)$)is unaffected by additional liquidity suppliers because the standard deviation of inventory is adjusted by the reciprocal of the adjustment to the conditional price pressure. The model-implied primitive parameters and the analysis of the spread (as summarized in Panel B and C of Table 4) are robust to this bias except for the private value rate λ which has to be adjusted by the reciprocal of the adjustment to the conditional price pressure.

4 Conclusion

Empirically we use 12 years of NYSE intermediary data to estimate price pressure—the deviation of prices from fundamental values due to the inventory risks born by an intermediary providing liquidity to asynchronously arriving investors. We construct a theoretical model to understand and characterize the effects of price pressure. The structure of the model allows for estimation of the intermediary' risk aversion and social costs of price pressure. We find:

1. A \$100,000 inventory shock causes price pressure of 1.01% for the small-capitalization stocks and 0.02% for the large-cap stocks.

2. The daily transitory volatility in stock returns due to price pressure (a measure of average price pressure) is large: 1.20% and 0.17% for small and large stocks, respectively. For small stocks the ratio of transitory volatility due to price pressure to the permanent (random-walk or efficient price) volatility is greater than one.

3. The model together with the time series properties of price pressure identifies low risk aversion for the intermediaries, a 0.10 coefficient of relative risk aversion, and deviations from constrained Pareto efficiency of 0.35 basis points of the value traded.

The significant size of price pressure suggest that a goal of financial market regulation should be to mitigate price pressure (see also SEC (2010)). One way to do this is by increasing capital for intermediation as the greater the risk bearing capacity of the intermediaries the smaller the price pressure. Another approach would be to lower costs for investors to monitor the market. This would lead to investor trading being more responsive to price pressures, reducing the duration of price pressure by allowing intermediaries to mean-revert their inventories more quickly.

Appendix I: Details on the likelihood optimization

The likelihood of the state space model described by equations (3-5) is optimized in the three below steps to minimize the probability of finding a local maximum. The optimization 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)).

1. An OLS regression of the log price series first difference on contemporaneous and lagged $\hat{f}_t$ yields starting values for β_i and $\beta_i^0, \ldots, \beta_i^k$ (see equations (3)-(4)). 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 $(\sigma(\varepsilon), \varphi)$ which are fixed at values that are picked from a nine by nine grid. φ ranges from 0 to 0.8 and $\sigma(\varepsilon)$ 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 9*9=81 optimizations and the $(\sigma(\varepsilon), \varphi)$ 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 one and a price pressure volatility that approaches the stock's unconditional volatility, i.e., it 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., $s_0 \sim \mathrm{N}(0, \frac{\sigma^2(\varepsilon)}{1-\varphi^2})$.
3. The likelihood is optimized where all parameters are free and starting values for $(\beta_i, \beta_i^0, \ldots, \beta_i^k, \sigma(\varepsilon), \varphi)$ 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 w.r.t. the parameters and (ii) the one-step change in parameter values (they both become arbitrarily small).

Appendix II

This appendix identifies the primitive parameters as functions of the empirical results documented in Tables 1, 2, and 3. They are identified by matching the observed time series pattern of price pressure ('$\alpha_i I_{it}$') of Section 1 to the stationary distribution of price pressure ('αi') implied by the model's closed-form solution of Section 2. In addition, trading volume is needed to separate the λ and θ parameters which characterize investors' demand for liquidity. To do this we must incorporate the fact that not all investor trading occurs via the intermediary. To correct for this we use the volume intermediated by the specialist, denote ν in the equations below, which is the fraction of trading volume in which the specialist is a participant, *specialist_particip*, times the total trading volume, *dollar_volume*. The effective (half-)spread is denoted δ^e. The expressions for the model's primitive parameters are:

$$\lambda = \frac{(-\alpha * \nu + \delta^e * (1 - \rho_I))^2}{2 * (-\alpha) * (1 - \rho_I)} \tag{26}$$

where ρ_I denotes the first-order autoregressive coefficient for inventory,

$$\frac{1}{\theta} = \frac{-\alpha * \nu + \delta^e * (1 - \rho_I)}{1 - \rho_I} \tag{27}$$

$$\gamma = \frac{2 * (\frac{1}{\beta} - \rho_I) * \alpha^2 * (1 - \rho_I^2) * \sigma^2(I)}{\rho_I^2 * \sigma^2(w)} \tag{28}$$

The analysis of the spread uses:

$$comp\_spread = \frac{-\alpha * (1 - \rho_I^2) * \sigma^2(I)}{\rho_I * \nu} \tag{29}$$

$$subsidy = \frac{-\alpha * \beta * (1 - \rho_I) * (1 - \rho_I^2) * \sigma^2(I)}{2 * (1 - \beta * \rho_I^2) * \nu} \tag{30}$$

$$pareto\_eff\_spread = \frac{-\alpha * (-(1 - \beta) + 2\beta RQ + \sqrt{(1 - \beta)^2 + 4 * \beta * (1 + \beta) * RQ + 4 * \beta^2 * (RQ)^2}) * (1 - \rho_I^2) * \sigma^2(I)}{4 * (1 - \rho_I) * \nu} \tag{31}$$

where

$$RQ \equiv (\frac{1}{\beta * \rho_I} - 1) * (1 - \rho_I) \tag{32}$$

Appendix III: Value of price pressure discount enjoyed by liquidity demander

Let w_i be the expected value of the price pressure discount enjoyed by the liquidity demander when the intermediary starts on an inventory of i. The Markovian law of motion for the intermediary's inventory position allows for Bellmanizing this value as follows:

$$w_i = \mathrm{E}_i[\frac{1}{2}Qs^2 + \beta w_{i'}] \tag{33}$$

Assume $w_i = A + Bi^2$ and calculate w_i from equation (33):

$$A + Bi^2 = w_i = \frac{1}{2}Q\alpha^2 i^2 + \beta \mathrm{E}_i[A + B(i + Q\alpha i + \varepsilon)^2] = \big(\beta A + \beta B\sigma_\varepsilon^2\big) + \big(\frac{1}{2}Q\alpha^2 + \beta B(1 + Q\alpha)^2\big)i^2 \tag{34}$$

A and B are solved by matching the constant and the coefficient of i^2 on both sides of the equation.

References

Amihud, Y. and H. Mendelson (1980). Dealership market: Market-making with inventory. *Journal of Financial Economics 8*, 31–53.

Barsky, R., T. Juster, M. Kimball, and M. Shapiro (1997). Preference parameters and behavioral heterogeneity: An experimental approach in the heath and retirement survey. *Quarterly Journal of Economics 112*, 537–579.

Brunnermeier, M. K. and L. H. Pedersen (2008). Market liquidity and funding liquidity. *Review of Financial Studies (forthcoming).*

Campbell, J. Y., S. J. Grossman, and J. Wang (1993). Trading volume and serial correlation in stocks returns. *The Quarterly Journal of Economics 108*, 905–939.

Campbell, J. Y., A. W. Lo, and A. C. MacKinlay (1997). *The Econometrics of Financial Markets.* Princeton, New Jersey: Princeton University Press.

Chetty, R. (2006). A new method of estimating risk aversion. *American Economic Review 96*, 1821–1834.

Cochrane, J. H. (2001). *Asset Pricing.* Princeton: Princeton University Press.

Cohen, A. and L. Einav (2007). Estimating risk preferences from deductible choice. *American Economic Review 97*, 745–788.

Coval, J. and E. Stafford (2007). Asset fire sales (and purchases) in equity markets. *Journal of Financial Economics 86*, 479–512.

Demsetz, H. (1968). The cost of transacting. *Quarterly Journal of Economics 82*, 33–53.

Duffie, D., N. Gârleanu, and L. Pedersen (2005). Over-the-counter markets.

Durbin, J. and S. J. Koopman (2001). *Time Series Analysis by State Space Models*. Oxford: Oxford University Press.

Evans, M. and R. Lyons (2002). Orderflow and exchange rate dynamics. *Journal of Political Economy 110*, 170–180.

Gabaix, X., P. Gopikrishnan, V. Plerou, and H. Stanley (2006). Institutional investors and stock market volatility. *Quarterly Journal of Economics 121*, 461–504.

Garman, M. B. (1976). Market microstructure. *Journal of Financial Economics 3*, 257–275.

Gavazza, A. (2009). The role of trading frictions in real asset markets. *American Economic Review (forthcoming)*.

Greenwood, R. (2005). Short- and long-term demand curves for stocks: Theory and evidence on the dynamics of arbitrage. *Journal of Financial Economics 75*, 607–649.

Grossman, S. J. and M. H. Miller (1988). Liquidity and market structure. *Journal of Finance 43*, 617–633.

Harris, L. and E. Gurel (1986). Price and volume effects associated with changes in the s&p 500 list: New evidence for the existence of price pressures. *Journal of Finance 41*, 815–829.

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

Hasbrouck, J. and G. Sofianos (1993). The trades of market makers: An empirical analysis of nyse specialists. *Journal of Finance 48*, 1565–1593.

Hendershott, T. and P. C. Moulton (2007). Speed and stock market quality: The nyses hybrid. Manuscript, UC Berkeley.

Hendershott, T. and M. S. Seasholes (2007). Market maker inventories and stock prices. *American Economic Review 97*, 210–214.

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

Huggett, M. (1993). The risk-free rate in heterogeneous agent incomplete insurance economies. *Journal of Economic Dynamics and Control 17*, 953–969.

Koopman, S., N. Shephard, and J. Doornik (1999). Statistical algorithms for models in state space using ssfpack 2.2. *Econometrics Journal 2*, 113–166.

Kraus, A. and H. Stoll (1972). Price impacts of block trading on the new york stock exchange. *Journal of Finance 27*, 569–588.

Lagos, R., G. Rocheteau, and P.-O. Weill (2009). Crashes and recoveries in illiquid markets. Manuscript, UCLA.

Levine, R. and S. Zervos (1998). Stock markets, banks, and economic growth. *American Economic Review 88*, 537–558.

Ljungqvist, L. and T. J. Sargent (2004). *Recursive Macroeconomic Theory*. Cambridge Massachusetts: The MIT Press.

Madhavan, A. and S. Smidt (1991). A bayesian model of intraday specialist pricing. *Journal of Financial Economics 30*, 99–134.

Madhavan, A. and S. Smidt (1993). An analysis of changes in specialist inventories and quotations. *Journal of Finance 48*, 1595–1628.

Mehra, R. and E. Prescott (1985). The equity premium: A puzzle. *Journal of Monetary Economics 15*, 145–161.

Menkveld, A. J., S. J. Koopman, and A. Lucas (2007). Modelling round-the-clock price discovery for cross-listed stocks using state space methods. *Journal of Business & Economic Statistics 25*, 213–225.

Mildenstein, E. and H. Schleef (1983). The optimal pricing policy of a monopolistic market maker in the equity market. *Journal of Finance 38*, 218–231.

Panayides, M. (2007). Affirmative obligations and market making with inventory. *Journal of Financial Economics 86*, 513–542.

Pastor, L. and R. Stambaugh (2003). Liquidity risk and expected returns. *Journal of Political Economy 111(3)*, 642–685.

Rock, K. (1990). The specialist's order book and price anomalies. Manuscript, Harvard University.

Rust, J. and J. Hall (2003). Middlemen versus market makers: A theory of competitive exchange. *Journal of Political Economy 111*, 353–403.

SEC (2010). Concept Release on Equity Market Structure, Release No. 34-61358; File No. S7-02-10.

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

Townsend, R. M. (1978). Intermediation with costly bilateral exchange. *Review of Economic Studies 45*, 417–425.

Weill, P.-O. (2007). Leaning against the wind. *Review of Economic Studies 74*, 1329–1354.

Zabel, E. (1981). Competitive price adjustment without market clearing. *Econometrica 49*, 1201–1221.

Table 1: Summary statistics

This table presents summary statistics on the dataset, which combines CRSP, TAQ, and a proprietary NYSE dataset. It is a balanced panel that contains daily observations on 697 NYSE common stocks from January 1994 through December 2005. Stocks are sorted into quintiles based on market capitalization, where quintile 1 contains large-cap stocks.

variable	description (units)	source	mean Q1	mean Q2	mean Q3	mean Q4	mean Q5	st. dev. within[a]
$midquote_{it}$	closing midquote, div/split adjusted[b] ($)	NYSE	53.76	44.52	36.65	28.58	19.21	22.20
$invent\_shares_{it}$	specialist inventory at the close (1,000 shares)	NYSE	8.19	5.55	4.33	3.23	5.39	34.19
$invent\_dollar_{it}$	specialist inventory at the close[b] ($1,000)	NYSE/CRSP	412.65	168.75	129.48	75.44	77.90	1,383.43
$shares\_outst_{it}$	shares outstanding (million)	CRSP	729.92	157.74	70.08	36.26	18.73	283.75
$market\_cap_{it}$	shares outstanding times price ($billion)	CRSP	34.29	5.34	2.06	0.88	0.29	11.57
$espread_{it}$	share-volume-weighted effective half spread (bps)	TAQ	8.41	12.46	16.50	24.60	46.12	24.20
$dollar\_volume_{it}$	average daily volume ($million)	TAQ	88.21	23.44	10.13	3.63	0.99	42.31
$specialist\_particip_{it}$	specialist participation rate (%)	NYSE	12.31	12.73	14.10	16.58	20.87	8.30
#observations: 697*3,018 (stock*day)								

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

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

Table 2: Price and inventory mean reversion estimates by year and size quintile

This table estimates the dynamics of log prices and intermediary inventories as well as their interaction. A common factor, $\hat{f}_t$ which is the cross-sectional mean each day of the standardized series, is removed from both series. For example, for each stock, each year we multiply specialist inventory in shares by the average price and then regress it on $\hat{f}_t$. The residuals from this regression are the idiosyncratic component of specialist inventory in dollars, I_{it}. For various dependent and independent variables we perform the following regressions by size quintile (Q1 contains the largest stocks) and by year:

$$y_{it} = \alpha_i + \beta_i x_{it} + \varepsilon_{it}$$

The table reports p-values in brackets. These p-values are based on a test statistic that counts the number of significant t-values across all stock-year estimates in the bin. The test statistic is binomially distributed under the null (we use the 0.1 and 0.9 quantiles in the t-test).

	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	all
Panel A: Autocorrelation 1st lag coef log price change ($x_{it} = y_{i,t-1}$)													
Q1	−0.03 (0.000)	−0.03 (0.000)	−0.05 (0.000)	−0.08 (0.000)	−0.02 (0.000)	−0.00 (0.005)	−0.01 (0.003)	−0.01 (0.001)	−0.05 (0.000)	−0.00 (0.005)	0.00 (0.005)	−0.02 (0.026)	−0.03 (0.000)
Q2	−0.02 (0.000)	−0.02 (0.000)	−0.03 (0.000)	−0.07 (0.000)	−0.01 (0.000)	−0.02 (0.000)	−0.04 (0.000)	−0.00 (0.000)	−0.06 (0.000)	−0.02 (0.000)	−0.02 (0.000)	−0.03 (0.000)	−0.03 (0.000)
Q3	−0.00 (0.000)	0.00 (0.000)	−0.02 (0.000)	−0.06 (0.000)	−0.01 (0.000)	−0.04 (0.000)	−0.06 (0.000)	−0.02 (0.000)	−0.06 (0.000)	−0.04 (0.000)	−0.03 (0.000)	−0.03 (0.001)	−0.03 (0.000)
Q4	−0.01 (0.000)	−0.02 (0.000)	−0.03 (0.000)	−0.05 (0.000)	−0.00 (0.000)	−0.03 (0.000)	−0.07 (0.000)	−0.04 (0.000)	−0.08 (0.000)	−0.06 (0.000)	−0.08 (0.000)	−0.04 (0.000)	−0.04 (0.000)
Q5	−0.05 (0.000)	−0.04 (0.000)	−0.02 (0.000)	0.00 (0.000)	0.04 (0.000)	0.01 (0.000)	−0.03 (0.000)	0.00 (0.000)	−0.02 (0.000)	−0.01 (0.000)	−0.04 (0.000)	−0.02 (0.000)	−0.02 (0.000)
all	−0.02 (0.000)	−0.02 (0.000)	−0.03 (0.000)	−0.05 (0.000)	−0.00 (0.000)	−0.02 (0.000)	−0.04 (0.000)	−0.01 (0.000)	−0.05 (0.000)	−0.03 (0.000)	−0.03 (0.000)	−0.03 (0.000)	−0.03 (0.000)
Panel B: Autocorrelation 2nd lag coef log price change ($x_{it} = y_{i,t-2}$)													
Q1	−0.03 (0.001)	−0.05 (0.000)	−0.04 (0.000)	−0.02 (0.094)	−0.03 (0.135)	−0.02 (0.001)	−0.06 (0.000)	−0.05 (0.000)	−0.01 (0.026)	−0.01 (0.249)	−0.02 (0.015)	−0.01 (0.135)	−0.03 (0.000)
Q2	−0.03 (0.001)	−0.03 (0.003)	−0.03 (0.009)	−0.01 (0.320)	−0.03 (0.001)	−0.01 (0.249)	−0.04 (0.001)	−0.02 (0.000)	−0.01 (0.015)	−0.02 (0.001)	−0.00 (0.009)	−0.02 (0.649)	−0.02 (0.000)
Q3	−0.02 (0.584)	−0.02 (0.070)	−0.02 (0.045)	−0.01 (0.664)	−0.01 (0.584)	−0.01 (0.029)	−0.04 (0.000)	−0.02 (0.017)	−0.00 (0.010)	−0.00 (0.029)	−0.01 (0.029)	−0.01 (0.001)	−0.02 (0.000)
Q4	−0.01 (0.001)	−0.02 (0.249)	−0.02 (0.041)	−0.03 (0.041)	−0.02 (0.320)	−0.00 (0.063)	−0.03 (0.009)	−0.02 (0.094)	0.00 (0.003)	−0.01 (0.003)	0.01 (0.003)	−0.00 (0.001)	−0.01 (0.000)
Q5	−0.01 (0.029)	−0.01 (0.000)	−0.01 (0.017)	−0.00 (0.017)	−0.00 (0.000)	0.00 (0.010)	−0.01 (0.500)	−0.00 (0.000)	−0.01 (0.017)	−0.01 (0.009)	0.00 (0.087)	0.00 (0.070)	−0.01 (0.000)
all	−0.02 (0.000)	−0.02 (0.000)	−0.02 (0.000)	−0.01 (0.010)	−0.02 (0.000)	−0.01 (0.000)	−0.04 (0.000)	−0.02 (0.000)	−0.00 (0.000)	−0.01 (0.000)	−0.00 (0.000)	−0.01 (0.000)	−0.02 (0.000)
Panel C: Standard deviation of idiosyncratic component specialist inventory I_{it}													
Q1	691 (0.000)	968 (0.000)	813 (0.000)	964 (0.000)	1126 (0.000)	1336 (0.000)	1344 (0.000)	1489 (0.000)	1472 (0.000)	1122 (0.000)	1119 (0.000)	1128 (0.000)	1131 (0.000)
Q2	472 (0.000)	510 (0.000)	488 (0.000)	524 (0.000)	530 (0.000)	695 (0.000)	819 (0.000)	647 (0.000)	448 (0.000)	391 (0.000)	441 (0.000)	400 (0.000)	530 (0.000)
Q3	374 (0.000)	429 (0.000)	383 (0.000)	372 (0.000)	430 (0.000)	452 (0.000)	668 (0.000)	437 (0.000)	293 (0.000)	242 (0.000)	266 (0.000)	271 (0.000)	385 (0.000)
Q4	226 (0.000)	254 (0.000)	255 (0.000)	261 (0.000)	291 (0.000)	315 (0.000)	320 (0.000)	333 (0.000)	229 (0.000)	163 (0.000)	145 (0.000)	147 (0.000)	245 (0.000)
Q5	167 (0.000)	159 (0.000)	167 (0.000)	234 (0.000)	204 (0.000)	223 (0.000)	210 (0.000)	186 (0.000)	129 (0.000)	108 (0.000)	95 (0.000)	95 (0.000)	165 (0.000)
all	386 (0.000)	464 (0.000)	421 (0.000)	471 (0.000)	516 (0.000)	604 (0.000)	672 (0.000)	619 (0.000)	514 (0.000)	405 (0.000)	413 (0.000)	408 (0.000)	491 (0.000)
Panel D: AR coef estimates idiosyncratic component specialist inventory I_{it} ($x_{it} = y_{i,t-1}$)													
Q1	0.28 (0.000)	0.27 (0.000)	0.26 (0.000)	0.22 (0.000)	0.25 (0.000)	0.27 (0.000)	0.28 (0.000)	0.29 (0.000)	0.28 (0.000)	0.34 (0.000)	0.37 (0.000)	0.25 (0.000)	0.28 (0.000)
Q2	0.47 (0.000)	0.46 (0.000)	0.44 (0.000)	0.38 (0.000)	0.35 (0.000)	0.34 (0.000)	0.36 (0.000)	0.32 (0.000)	0.25 (0.000)	0.28 (0.000)	0.33 (0.000)	0.25 (0.000)	0.35 (0.000)
Q3	0.59 (0.000)	0.59 (0.000)	0.56 (0.000)	0.51 (0.000)	0.49 (0.000)	0.45 (0.000)	0.41 (0.000)	0.41 (0.000)	0.30 (0.000)	0.31 (0.000)	0.34 (0.000)	0.24 (0.000)	0.43 (0.000)
Q4	0.74 (0.000)	0.73 (0.000)	0.71 (0.000)	0.66 (0.000)	0.63 (0.000)	0.63 (0.000)	0.59 (0.000)	0.57 (0.000)	0.40 (0.000)	0.38 (0.000)	0.36 (0.000)	0.29 (0.000)	0.56 (0.000)
Q5	0.82 (0.000)	0.80 (0.000)	0.80 (0.000)	0.77 (0.000)	0.78 (0.000)	0.79 (0.000)	0.77 (0.000)	0.76 (0.000)	0.66 (0.000)	0.61 (0.000)	0.57 (0.000)	0.51 (0.000)	0.72 (0.000)
all	0.58 (0.000)	0.57 (0.000)	0.56 (0.000)	0.51 (0.000)	0.50 (0.000)	0.50 (0.000)	0.48 (0.000)	0.47 (0.000)	0.38 (0.000)	0.38 (0.000)	0.39 (0.000)	0.31 (0.000)	0.47 (0.000)

<continued on next page>

<continued from previous page>

	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	all
Panel E: Regression coefficient log price change on lagged idiosyncratic component specialist inventory $I_{i,t-1}$													
Q1	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.01
	(0.005)	(0.000)	(0.009)	(0.000)	(0.005)	(0.135)	(0.005)	(0.005)	(0.001)	(0.249)	(0.063)	(0.135)	(0.000)
Q2	0.01	0.01	0.01	0.02	0.01	0.02	0.02	0.02	0.02	0.02	0.01	0.01	0.02
	(0.041)	(0.009)	(0.015)	(0.000)	(0.041)	(0.000)	(0.001)	(0.041)	(0.000)	(0.135)	(0.135)	(0.063)	(0.000)
Q3	0.02	0.02	0.02	0.02	0.02	0.03	0.05	0.01	0.05	0.03	0.03	0.03	0.03
	(0.010)	(0.001)	(0.045)	(0.000)	(0.145)	(0.001)	(0.000)	(0.000)	(0.000)	(0.017)	(0.102)	(0.199)	(0.000)
Q4	0.03	0.02	0.04	0.05	0.04	0.06	0.08	0.05	0.11	0.05	0.08	0.06	0.06
	(0.063)	(0.135)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.026)	(0.000)	(0.026)	(0.000)
Q5	0.08	0.08	0.03	0.02	0.06	0.06	0.14	0.10	0.15	0.08	0.12	0.13	0.09
	(0.000)	(0.199)	(0.010)	(0.017)	(0.006)	(0.416)	(0.000)	(0.000)	(0.000)	(0.009)	(0.008)	(0.500)	(0.000)
all	0.03	0.03	0.02	0.02	0.03	0.03	0.06	0.04	0.07	0.04	0.05	0.05	0.04
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.008)	(0.000)

*/**: Significant at a 95%/99% level.

Table 3: State space model estimates by year and size quintile

This table estimates the following state space model for a latent efficient price and an observed end-of-day midquote price:

$$
\begin{aligned}
\text{(observed price)} \quad p_{it} &= m_{it} + s_{it} \\
\text{(unobserved efficient price)} \quad m_{it} &= m_{i,t-1} + \beta_i \hat{f}_t + w_{it} \quad w_{it} = \kappa_i \hat{I}_{it} + u_{it} \\
\text{(unobserved transitory price deviation)} \quad s_{it} &= \alpha_i I_{it} + \beta_i^0 \hat{f}_t + \cdots + \beta_i^3 \hat{f}_{t-3} + \varepsilon_{it}
\end{aligned}
$$

where i indexes over stocks and t indexes over days, m_{it} is the end-of-day unobserved efficient price ('state'), $\hat{f}_t$ is a midquote return common factor which is the cross-sectional average of the standardized midquote return series which has been filtered with an AR(4) model to remove intertemporal dynamics, p_{it} is end-of-day observed midquote, I_{it} is the idiosyncratic part of the specialist end-of-day inventory in dollars that remains after removing a common factor across specialist inventories, $\hat{I}_{it}$ is the idiosyncratic inventory innovation which is obtained as the residual of an AR(9) model and captures the surprise part of the net imbalance, β_i^j captures potential 'overreaction' or lagged adjustment to common factor innovations, and w_{it} and ε_{it} are mutually independent i.i.d. error terms. The model is estimated using maximum likelihood estimates where the error terms w_{it} and ε_{it} are assumed to be normally distributed. The optimization is implemented in ox with ssfpack routines where we use the Kalman filter to evaluate the likelihood (see Koopman, Shephard, and Doornik (1999)). The table reports p-values in brackets. These p-values are based on a test statistic that counts the number of significant t-values across all stock-year estimates in the bin. The test statistic is binomially distributed under the null (we use the 0.1 and 0.9 quantiles in the t-test).

	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	all
Panel A: α_i conditional price pressure													
Q1	−0.03	−0.02	−0.02	−0.02	−0.02	−0.02	−0.02	−0.02	−0.02	−0.01	−0.01	−0.01	−0.02
	(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)
Q2	−0.06	−0.05	−0.05	−0.04	−0.03	−0.05	−0.05	−0.05	−0.05	−0.03	−0.03	−0.03	−0.04
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Q3	−0.11	−0.09	−0.09	−0.07	−0.09	−0.12	−0.11	−0.06	−0.09	−0.06	−0.07	−0.06	−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)
Q4	−0.33	−0.25	−0.30	−0.27	−0.26	−0.34	−0.27	−0.23	−0.27	−0.14	−0.19	−0.16	−0.25
	(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)
Q5	−1.07	−1.04	−0.78	−0.74	−1.01	−1.09	−1.19	−1.30	−1.20	−0.86	−0.94	−0.87	−1.01
	(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)
all	−0.32	−0.29	−0.25	−0.23	−0.28	−0.32	−0.33	−0.33	−0.32	−0.22	−0.25	−0.23	−0.28
	(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)
Panel B: $\lvert\alpha_i\rvert\sigma(I)_i$ (explained) transitory volatility i.e. unconditional price pressure													
Q1	14	13	13	16	19	22	25	30	22	15	14	9	17
Q2	26	24	22	19	20	33	36	30	23	14	15	13	23
Q3	39	40	35	31	40	48	45	35	23	16	20	17	32
Q4	65	64	69	58	63	86	77	63	41	21	24	21	54
Q5	116	109	115	107	152	177	184	155	108	92	71	54	120
all	52	50	51	46	58	73	73	62	43	31	29	23	49
Panel C: κ_i informativeness order imbalance innovation $\hat{I}$													
Q1	−0.08	−0.08	−0.08	−0.07	−0.08	−0.08	−0.09	−0.07	−0.07	−0.06	−0.06	−0.06	−0.07
	(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)
Q2	−0.16	−0.13	−0.13	−0.14	−0.17	−0.15	−0.18	−0.15	−0.16	−0.15	−0.13	−0.13	−0.15
	(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)
Q3	−0.22	−0.19	−0.18	−0.21	−0.25	−0.23	−0.26	−0.24	−0.26	−0.25	−0.26	−0.23	−0.23
	(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)
Q4	−0.34	−0.28	−0.29	−0.37	−0.44	−0.36	−0.47	−0.43	−0.52	−0.56	−0.52	−0.50	−0.43
	(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)
Q5	−0.59	−0.43	−0.61	−0.63	−0.72	−0.48	−0.73	−0.87	−1.25	−1.43	−1.29	−1.41	−0.87
	(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)
all	−0.28	−0.22	−0.26	−0.28	−0.34	−0.26	−0.35	−0.35	−0.46	−0.49	−0.45	−0.47	−0.35
	(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)
Panel D: $\lvert\kappa_i\rvert\sigma(\hat{I})_i$ explained permanent volatility													
Q1	39	36	38	44	57	65	80	56	55	37	30	27	47
Q2	42	39	41	45	65	70	85	61	56	44	36	36	52
Q3	44	42	43	54	72	70	86	75	60	48	41	36	56
Q4	47	40	42	52	73	64	83	75	66	60	49	43	58
Q5	43	39	49	56	69	56	78	90	80	75	65	63	64
all	43	39	43	50	67	65	83	72	63	53	44	41	55

<continued on next page>

<continued from previous page>

	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	all
Panel E: $\sigma(w)_i$ permanent volatility													
Q1	124	124	128	134	176	198	243	192	187	126	106	103	153
Q2	143	136	138	144	186	213	257	212	206	151	127	125	170
Q3	151	143	147	158	204	213	253	212	205	153	138	136	176
Q4	162	152	153	158	201	211	239	212	199	163	145	151	179
Q5	176	170	181	181	228	232	264	245	244	202	178	190	208
all	151	145	149	155	199	213	251	215	208	159	139	141	177
Panel F: $\frac{\alpha_i^2\sigma^2(I)_i}{\sigma^2(w)_i+\beta_i^2\sigma^2(\hat{f})}$ ratio of transitory and permanent variance													
Q1	0.02	0.02	0.01	0.02	0.01	0.02	0.01	0.03	0.02	0.02	0.03	0.01	0.02
Q2	0.08	0.09	0.04	0.03	0.02	0.04	0.03	0.03	0.02	0.01	0.02	0.02	0.03
Q3	0.28	0.27	0.16	0.09	0.13	0.19	0.05	0.07	0.02	0.01	0.04	0.09	0.12
Q4	0.30	0.47	0.47	0.28	0.23	0.47	0.22	0.16	0.06	0.03	0.04	0.05	0.23
Q5	1.11	1.42	1.14	1.16	1.24	2.55	1.96	1.09	0.79	2.10	0.61	0.62	1.32
all	0.36	0.45	0.36	0.31	0.32	0.65	0.46	0.28	0.18	0.44	0.15	0.16	0.34
Panel G: $\frac{\alpha_i^2\sigma^2(I)_i}{\sigma^2(w)_i}$ ratio of transitory and permanent 'idiosyncratic' variance													
Q1	0.03	0.03	0.02	0.03	0.02	0.02	0.02	0.04	0.03	0.04	0.04	0.01	0.03
Q2	0.10	0.10	0.05	0.04	0.02	0.05	0.04	0.04	0.03	0.02	0.03	0.02	0.04
Q3	0.35	0.28	0.20	0.11	0.18	0.21	0.06	0.10	0.04	0.02	0.05	0.12	0.14
Q4	0.36	0.48	0.54	0.31	0.29	0.51	0.25	0.23	0.08	0.04	0.05	0.05	0.27
Q5	1.28	1.48	1.26	1.28	1.51	2.64	2.16	1.40	0.98	2.31	0.73	0.67	1.47
all	0.42	0.47	0.41	0.35	0.40	0.68	0.51	0.36	0.23	0.48	0.18	0.18	0.39
Panel H: $\sigma(\varepsilon)_i$ error term transitory volatility													
Q1	20	17	25	30	24	20	31	22	31	18	12	15	22
	(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)
Q2	19	19	21	28	24	22	43	21	39	23	17	19	25
	(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)
Q3	19	17	22	23	27	30	45	27	41	25	20	20	26
	(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)
Q4	28	25	23	26	18	23	42	30	43	30	30	24	29
	(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)
Q5	42	42	34	25	21	29	47	38	39	28	28	25	33
	(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)
all	26	24	25	27	23	25	42	28	39	25	21	21	27
	(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)

Table 4: Model's primitive parameters and an analysis of the spread

This table analyzes the cost to liquidity demanders of a competitive intermediary who uses price pressure to control her inventory position. The model's primitive parameters consist of the liquidity demander's total private value rate λ and its dispersion $\frac{1}{\theta}$, uncertainty about net liquidity demand $\sigma(\varepsilon)$, a discount factor β, and the coefficient of relative risk aversion (γ) which is the model's absolute risk aversion ($\tilde{\gamma}$) divided by the intermediary's expected consumption at the competitive spread. The conditional price pressure estimates along with other trading variables measured in the empirical part of the paper (sample averages in Panel A cf. Table 1 and 3) allows for identification of the primitive parameters (Panel B) which in turn allow for a decomposition of the net cost to liquidity demanders (investors) of liquidity supply by a competitive intermediary (Panel C). It expresses this net cost as a fraction of transacted volume ('spread') and decomposes it into a spread received by the intermediary and a subsidy enjoyed by the liquidity demanders at times when prices are pressured. It also calculates the Pareto efficient cost if the intermediary were to internalize the subsidy enjoyed by the liquidity demander. Panel B and C report medians and interquartile ranges ($Q_{0.75}$-$Q_{0.25}$ where Q_i is quantile i) in parentheses.

	Q1	Q2	Q3	Q4	Q5	all
Panel A: Measured variables that identify model's primitive parameters (cf. Table 1 and 3)						
conditional price pressure α_i (bps per \$1000)	-0.02	-0.04	-0.09	-0.25	-1.01	-0.28
stdev daily inventory $\sigma(I)$ (\$1000)	1131	530	385	245	165	491
1st order autocorrelation inventory	0.28	0.35	0.43	0.56	0.72	0.47
price risk inventory $\sigma(w)_i$ (bps)	153	170	176	179	208	177
daily dollar volume[a] (\$million)	10.86	2.98	1.43	0.60	0.21	3.22
effective half spread	8.41	12.46	16.50	24.60	46.12	21.62
daily discount factor[b]	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
Panel B: Identification of model's primitive parameters						
daily private value rate λ (\$1000)	59.69 (241.59)	14.12 (37.71)	6.21 (13.39)	2.77 (4.63)	1.46 (2.31)	5.86 (21.73)
dispersion private value $\frac{1}{\theta}$ (bps)	171 (312)	135 (201)	148 (166)	187 (206)	307 (370)	186 (260)
intermediary's relative risk aversion γ	0.24 (1.06)	0.15 (0.55)	0.11 (0.35)	0.09 (0.26)	0.05 (0.13)	0.10 (0.35)
Panel C: Decomposition of the spread paid by liquidity demander						
(1) model-implied competitive spread (bps)	7.32 (15.81)	12.86 (25.31)	19.43 (40.04)	37.35 (68.62)	73.16 (149.27)	22.60 (57.68)
(2) price pressure subsidy to liq demander (bps)	0.58 (1.21)	1.40 (3.13)	2.63 (5.64)	6.00 (13.49)	15.52 (33.26)	2.88 (9.56)
(3) net spread to liq demander[c] (1)-(2) (bps)	6.50 (14.14)	11.01 (22.33)	16.52 (33.16)	30.33 (55.40)	57.41 (114.93)	19.14 (47.18)
(4) constrained Pareto efficient spread (bps)	6.45 (14.07)	10.74 (21.59)	16.05 (31.87)	29.54 (53.53)	54.77 (110.32)	18.67 (45.48)
(5) deadweight loss[c] (3)-(4) (bps)	0.05 (0.11)	0.14 (0.38)	0.30 (0.80)	0.88 (2.18)	2.58 (5.68)	0.35 (1.51)

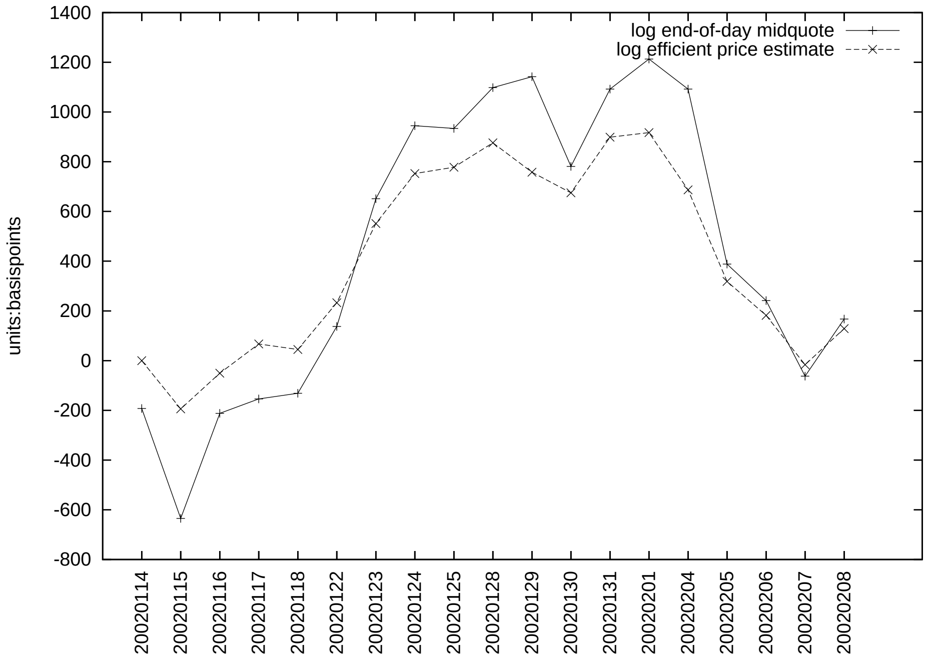
[a]: $dollar\_volume_{it} * specialist\_particip_{it}$ as a proxy for volume intermediated by the specialist
[b]: discount factor equals the reciprocal of the gross riskfree rate from Kenneth French' website
[c]: the reported difference is the median of differences, not the difference of (reported) medians

Figure 1: Price change decomposition given state space model estimates not using inventory data

This figure depicts the observed end-of-day log price series (midquote price) and its unobserved efficient price (given *all* data using the Kalman smoother) estimate not using inventory data for a representative stock (Rex Stores Corporation, ticker RSC, CRSP PERMNO 68830). The model is defined as:

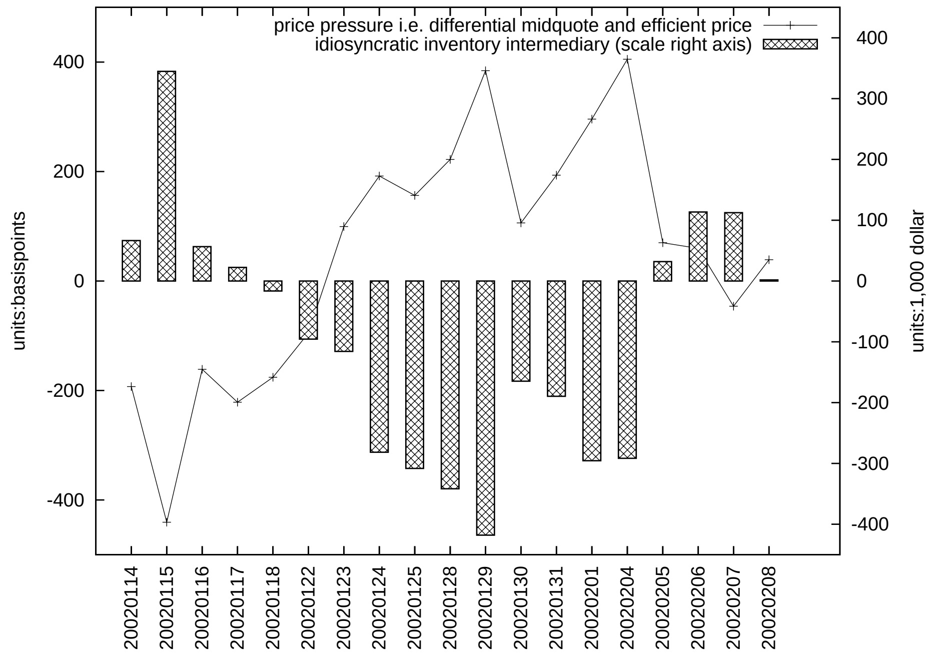
$$\begin{aligned} \text{(observed price)} \quad p_{it} &= m_{it} + s_{it} \\ \text{(unobserved efficient price)} \quad m_{it} &= m_{i,t-1} + \beta_i \hat{f}_t + w_{it} \\ \text{(unobserved transitory price deviation)} \quad s_{it} &= \varphi_i s_{i,t-1} + \beta_i^0 \hat{f}_t + \cdots + \beta_i^3 \hat{f}_{t-3} + \varepsilon_{it} \end{aligned}$$

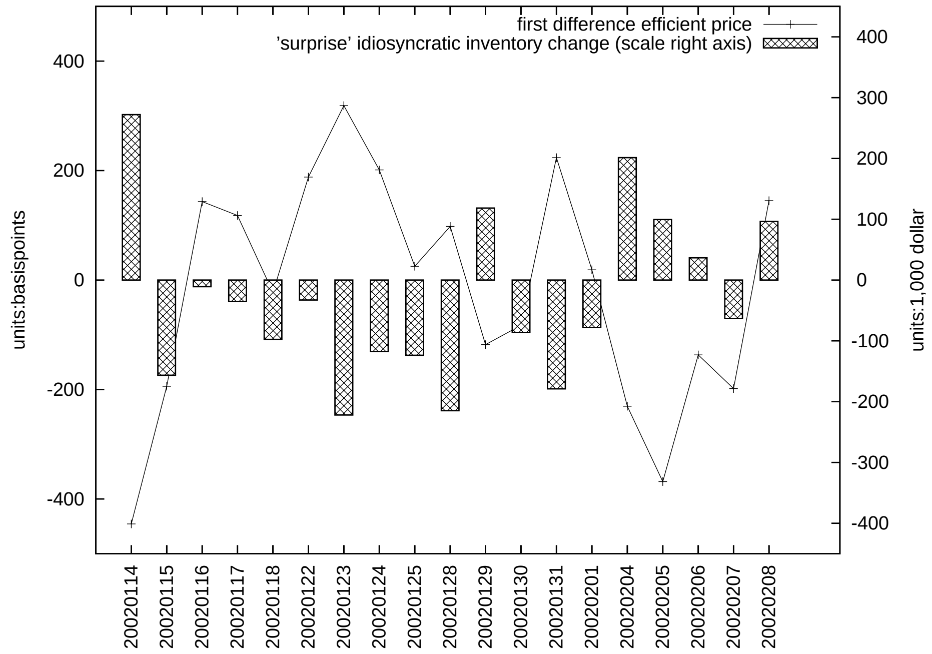
where i indexes over stocks and t indexes over days, m_{it} is the end-of-day unobserved efficient price ('state'), $\hat{f}_t$ is the midquote return common factor which is the cross-sectional average of the standardized midquote return series which has been filtered with an AR(4) model to remove intertemporal dynamics, p_{it} is the end-of-day observed midquote price, β_i^j captures potential 'overreaction' or lagged adjustment to common factor innovations, and w_{it} and ε_{it} are mutually independent i.i.d. error terms. The model is estimated using maximum likelihood estimates where the error terms w_{it} and ε_{it} are assumed to be normally distributed. The first panel graphs these series recentered around the first day's estimate of the efficient price. The second panel graphs the price pressure—the difference between the observed price and the efficient price—against the idiosyncratic inventory position of the specialist. The third graphs the efficient price innovation against the contemporaneous unpredictable 'surprise' idiosyncratic inventory change which is obtained as the residual of an AR(9) model.



<continued on next page>

<continued from previous page>





CFS Working Paper Series:

No.	Author(s)	Title
2010/13	**Sarah Draus**	Does Inter-Market Competition Lead to Less Regulation?
2010/12	**Kai-Oliver Maurer** **Carsten Schäfer**	Analysis of Binary Trading Patterns in Xetra
2010/11	**Annamaria Lusardi** **Olivia S. Mitchell**	How Ordinary Consumers Make Complex Economic Decisions: Financial Literacy and Retirement Readiness
2010/10	**Annamaria Lusardi** **Daniel Schneider** **Peter Tufano**	The Economic Crisis and Medical Care Usage
2010/09	**Annamaria Lusardi** **Olivia S. Mitchell** **Vilsa Curto**	Financial Literacy among the Young: Evidence and Implications for Consumer Polic
2010/08	**Volker Wieland** **Maik Wolters**	The Diversity of Forecasts from Macroeconomic Models of the U.S. Economy
2010/07	**Antje Brunner** **Jan Pieter Krahnen**	Hold-Up in Multiple Banking: Evidence from SME Lending
2010/06	**Roman Kräussl** **André Lucas** **David R. Rijsbergen,** **Pieter Jelle van der Sluis** **Evert B. Vrugt**	Washington Meets Wall Street: A Closer Examination of the Presidential Cycle Puzzle
2010/05	**Sander J.J. Konijn** **Roman Kräussl** **Andre Lucas**	Blockholder Dispersion and Firm Value

Copies of working papers can be downloaded at http://www.ifk-cfs.de