

Price Pressures¹

Terrence Hendershott and Albert J. Menkveld

current version: July 15, 2011
first version: November 12, 2009

(for supplementary material click <http://goo.gl/8ynea>)

¹Terrence Hendershott, Haas School of Business, University of California at Berkeley, 545 Student Services Bldg #1900, Berkeley, CA 94720, tel +1 510 643 0619, hender@haas.berkeley.edu. Albert J. Menkveld, VU University Amsterdam, FEWEB, De Boelelaan 1105, 1081 HV, Amsterdam, Netherlands, tel +31 20 598 6130, albertjmenkveld@gmail.com and TI Duisenberg School of Finance. We benefit from in-depth discussions with Viral Acharya, Bruno Biais, Sabrina Buti, Robert Engle, Joel Hasbrouck, Charles Jones, Jeremy Large, Sunny Li, Virgiliu Midrigan, Mark Seasholes, Stijn van Nieuwerburgh, Marios Panayides, Matthew Pritsker, Jan van Schuppen, Clara Vega, Pierre-Olivier Weill, and audience participation at the Bank of Canada, Cass, Dauphine, ESSEC, NBER microstructure, SOFIE/Chicago, Stanford, Swiss Central Bank, and U of Toronto. The project was inspired by the Ljungqvist and Sargent NYU PhD course 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.

Price Pressures

Abstract

We study price pressures — deviations from fundamental value due to risk-averse intermediaries supplying liquidity to asynchronously arriving investors. Empirically, New York Stock Exchange intermediary data reveal economically large price pressures, 0.49% on average with a half life of 0.92 days. Theoretically, a simple dynamic inventory model captures an intermediary's use of price pressure to mean-revert inventory. She trades off revenue loss due to price pressure against price risk associated with staying in a nonzero inventory state. A closed-form solution identifies the intermediary's risk aversion and the investors' private value distribution from the observed time series patterns. These parameters imply a relative social cost due to price pressure — a deviation from constrained Pareto efficiency — of approximately 10% of the cost of immediacy.

JEL: G12, G14, D53, D61

Keywords: liquidity, inventory risk, intermediary, volatility

For supplementary material click <http://goo.gl/8ynea>

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 twelve years of daily inventory data from intermediaries ('specialists') 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 and duration. These price pressures are the effective cost to investors who demand liquidity to trade large quantities quickly. We construct a recursive dynamic inventory model to identify intermediaries' risk aversion and investors' private value distribution.

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 \(1978\)](#)). Intermediaries offer 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 intermediaries' return if they simultaneously cross (execute 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

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

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.

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 NYSE intermediary inventory data. To link these inventory positions with price pressures we use the Kalman filter to estimate a state space model that decomposes stock prices into the sum of a martingale series and a stationary series. The martingale is best interpreted as a latent efficient price which is unpredictable and typically construed to track information. The stationary series identifies 'pricing errors' which the model allows to be directly related to the inventory of the intermediaries. The estimated coefficient in the stationary equation represents the transitory impact of a dollar of inventory, which we refer to as the conditional price pressure. The size of inventory shocks and their duration determine the (unconditional) volatility of inventory. Combining the conditional price pressure with inventory volatility yields the average size of price pressure which we refer to as the transitory volatility due to price 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 0.98 basis points of price pressure. The small conditional price pressure in large stocks is associated with large inventory deviations: 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 intermediaries' inventory is 0.55 days for the largest stocks and 2.11 days for the smallest stocks.

³Kraus 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) study 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 (2009) provide a theoretical model where price pressures affect intermediary capital positions due to margin requirements. This can require the intermediary to reduce leverage causing feedback effects in price pressures.

⁴Madhavan 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.

⁵Throughout 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.

Combining these results yield estimates of the daily transitory volatility due to price pressure that range from 0.17% for the largest stocks to 1.18% 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 an 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. In doing this the intermediary bears the cost of setting the average of the bid and ask quotes — the midquote price — 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 contribution of our theoretical approach is to facilitate estimation of the intermediary's risk aversion, the dispersion of investors' private values, and social welfare lost due to deviations from constrained Pareto efficiency. In the model the intermediary's first-order condition equates the size of the pecuniary subsidy, i.e., price pressure, to revert out of a nonzero position to the disutility of price risk associated with the nonzero position. The size of the price pressure identifies the intermediary's risk aversion. Private value dispersion is identified by how much position reversion is accelerated by an additional dollar of price pressure, i.e., it measures the elasticity of the relative amount of investor buying and selling. 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, private values uniformly-distributed with a range from -185 to +185 basis points. Finally, the price-setting intermediary does not fully internalize the costs and benefits of using price pressure to manage risk. The intermediary subsidizes investor demand for immediacy to reduce risk, but does not internalize the benefit to investors. We estimate the less than socially optimal amount of price pressure to be approximately 10% of total liquidity/immediacy costs.

The low risk aversion could arise from risk tolerant capital migrating to the intermediation sector or be a sign

of an agency conflict between the 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. The risk aversion might also be underestimated as the analysis assumes the intermediary cannot hedge any of the non-systematic risk on her position. In practice hedging may be possible, for example, by taking an offsetting position in stocks in the same industry.

Section 1 empirically estimates the daily magnitude of price pressures and pricing errors. Section 2 sets up and solves a structural model enabling use of the price pressure estimates to identify deeper parameters of the trading process. How the unique NYSE market structure and the intermediaries' role in it impact the interpretation of the results is discussed in Section 3. In particular, we evaluate an alternative, information-based explanation (cream skimming) for the negative correlation of inventories with pricing errors.

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 assigned to each stock one intermediary, called a specialist, to act as a market maker. This structure helps identify the amount of liquidity supplied by intermediaries. 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 price pressure and intermediaries' inventory risk exist in all markets where investor trading needs are not perfectly synchronized. Section 3 further discusses the NYSE structure and how it changed subsequent to our sample period.

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) data, and a proprietary NYSE dataset to construct the end-of-day midquote price (i.e., the average of the bid price and ask price), the end-of-day specialist inventory position, and 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 positions, which are in shares in the original database, into dollars by multiplying the position with 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]

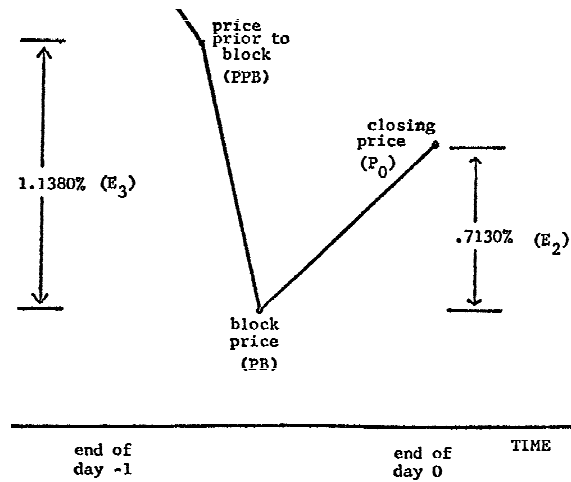
Summary statistics. Table 1 presents the mean of various trading variables by size quintile. It includes the within variation (defined as the standard deviation of the data series after removing stock fixed effects) to provide a sense of the variable's variability through time. Several observations emerge 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 specialists are active intermediaries 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 intermediaries is positive and economically significant.⁷ For example, for the large-cap stocks in Q1 they maintain 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.

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

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

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 intermediaries' 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 that prices overshoot in the event of a block trade, i.e., a transaction where the initiating party wants to trade a 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 (pricing error) that captures transitory price effects including 'price pressures'. We use log prices throughout the paper. In its simplest form the model structure for the log price is:

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

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

where p_t is the observed price, m_t is the unobserved efficient price, s_t is the unobserved pricing error, δ_t is the required return obtained as the (time-varying) riskfree rate plus beta times a 6% risk premium, and w_t is the innovation in the efficient price. s_t and w_t are assumed to be 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 since it affects prices only at a single point in time. The model can be estimated with maximum likelihood where the likelihood is constructed using 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 (Δp_t) is an invertible first-order moving average (MA(1)) time series model which implies an 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) has 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 *all* observations, i.e., past prices, the current price, 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 we focus on.

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:

$$m_{it} = m_{i,t-1} + \delta_{it} + \beta_i f_t + w_{it}, \quad w_{it} = \kappa_i \hat{I}_{it} + u_{it} \quad (3)$$

where the subscript i indexes stocks, t indexes days, δ_t is the required return obtained as the (time-varying) riskfree rate plus beta times a 6% risk premium, f_t is the demeaned market return (CRSP value-weighted return), w_{it} is the idiosyncratic innovation, \hat{I}_{it} is the idiosyncratic inventory innovation (obtained as the residual of an nine lag autoregressive model model⁸ (AR(9)) applied to market-cap weighted standardized inventory) 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.⁹ The decomposition of the efficient price innovation into a common factor component ($\beta_i f_t$) 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 again easily hedged through index products.

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

$$s_{it} = \alpha_i I_{it} + \beta_i^0 f_t + \dots + \beta_i^k f_{t-k} + \varepsilon_{it} \quad (4)$$

where the error term ε_{it} is normally distributed and uncorrelated with u_{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 pricing error equation.

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

$$p_{it} = m_{it} + s_{it}. \quad (5)$$

⁸Throughout the empirical analysis standard information criteria are used to determine the number of lags in an autoregressive model.

⁹While 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 the pricing error (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).

As we intend to analyze price pressures in the cross-section as well as in the time dimension all empirical analyses are done by stock-year. To report the 697*12 stock-year results we aggregate stocks into quintile bins according to their size. This allocation of stocks to size-quintiles is based on the full-sample so, for example, the large-cap quintile therefore contains the same set of stocks in each year of the sample. Means are calculated for each size-year bin separately and by also by pooling all size-bins across years. These quintile-averages are reported in the ‘all’ column of the tables and t -statistics for these are based on stock and year clustered standard errors (cf. [Petersen \(2009\)](#)).

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 transitory price effects that we are after are present in the data.

[insert Table 2]

Panels A and B of Table 2 report 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 stock autocorrelation results in [Campbell, Lo, and MacKinlay \(1997\)](#) the average first order autocorrelation is negative in 65 of the 70 size-year bins. The t -statistics calculated by quintile and across all quantiles are all significant at conventional levels. The negative first order autocorrelation is consistent with pricing errors 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 only an 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.¹⁰

¹⁰If 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 is (the higher φ is), the less negative the return autocorrelation becomes.

Panels C, D, and E of Table 2 report the standard deviation of intermediaries' 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 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 intermediaries will shy away from frequent and large nonzero inventory positions, e.g., see the model 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 shows 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.¹¹ 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 takes multiple days for the small-cap stocks. 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 results of 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 the pricing error s_t to the intermediary's inventory series. Initially, we do

¹¹These 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).

not use inventory data to estimate a reduced version of the model:

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

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

$$s_{it} = \varphi_i s_{i,t-1} + \beta_i^0 f_t + \dots + \beta_i^3 f_{t-3} + \varepsilon_{it} \tag{8}$$

where the AR(1) process for s_{it} allows pricing errors 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 pricing error processes conditional on *all* observations. In other words, it uses past prices, the current price, 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. We next test these observations 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. Prices are temporarily low when the intermediary is on a long position and temporarily 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 -0.98 for the small-cap stocks. These numbers are economically significant even in the large stocks as a \$1.131 million (one standard deviation) position change in intermediary inventory creates a price pressure of $1131 \times 0.02 = 17$ basis points (cf. effective half-spread of 8 basis points, see Table 1).

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 118 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 1.28, 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 -0.34 on average and statistically significant. In Panel D the average explained permanent volatility — κ_i times the standard deviation of unexpected inventory changes — is 54 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 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 (including the transition probabilities) 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.

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. 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 maximizes quadratic utility over nonstorable consumption,¹² and (iii) a security position exposes the intermediary to fundamental value risk which is modeled as a normally distributed stochastic dividend¹³ to avoid a notational burden.¹⁴

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

¹²[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.

¹³The stochastic dividend minimizes accounting in the model, but, as [Ho and Stoll \(1981, p.52\)](#) emphasize, no sources of uncertainty are ignored.

¹⁴These 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.

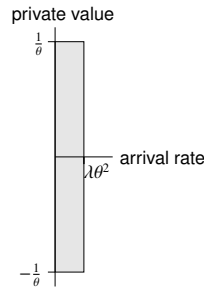
asset to outside investors who demand liquidity. At the start of each period a single intermediary quotes an ask price at which she commits to sell and a bid price at which she commits to buy.

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 the bid-ask spread of 2δ and price pressure of s , analogous to $s_t = p_t - m_t$ in the state space model. Specifically, we assume investor buying and selling demand for liquidity are normally distributed variables linear in the bid-ask quotes:¹⁵

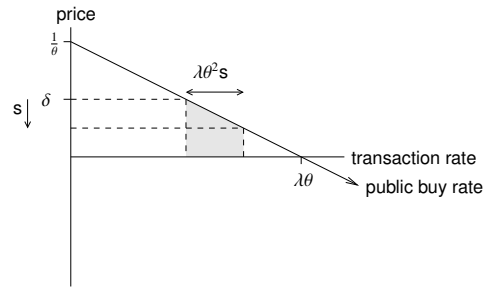
$$q_s(s, \delta) = \lambda\theta(1 - \theta(s + \delta)) + \varepsilon_s, \quad \varepsilon_s \sim \mathcal{N}\left(0, \frac{1}{2}\sigma_\varepsilon^2\right), \quad (9)$$

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

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.



(i) mean demand per private value



(ii) mean public buy volume as function of ask price

¹⁵These distributions can arise from an economy where every time period each of a finite number N of investors 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 investors' hedging or liquidity needs from which they derive private values from trading now versus when they next contact the market (see Grossman and Miller (1988) and Duffie (2010, sections III and IV)). The distribution of private values arises from the investors' intertemporal optimization problem across their current and future contacts with the market (see Duffie, Gârleanu, and Pedersen (2005) and Weill (2007) for models directly incorporating this optimization). If private value shocks are uniformly distributed on support $[-\frac{1}{\theta}, \frac{1}{\theta}]$ and 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 symmetrically distributed.

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-horizon dynamic program:

$$v_{i_0} = \max_{\{s_t(i'), \delta_t(i')\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} E_{i_0} \left(\beta^t \left((c_t - \bar{c}) - \frac{1}{2} \tilde{\gamma} (c_t - \bar{c})^2 \right) \right) \quad (11)$$

where i' represents the history of her inventory position through time t and $E_{i_0}(\cdot)$ 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 \bar{c} which is fixed ex-ante at the competitive intermediary's average consumption.¹⁶ The intermediary's actual consumption in period t equals net trading revenue plus the stochastic dividend:

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

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

$$E_t(c_t) = 2\lambda\theta(\delta_t - \theta(s_t^2 + \delta_t^2)) \quad (13)$$

$$\text{var}_t(c_t) = \sigma_\varepsilon^2(s_t^2 + \delta_t^2) + \sigma^2 i_t^2 \quad (14)$$

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

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

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 conditional mean. This removes the quadratic difference between the conditional mean and \bar{c} (the unconditional mean), which would make the objective function fourth order and not possible 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

¹⁶ \bar{c} is assumed to be exogenous for the optimizing agent. Doing this makes for convenient calculation of the intermediary's relative risk aversion later in the section.

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:

$$v_{i_0} = \max_{\{s_t(i'), \delta_t(i')\}_{i=0}^\infty} \sum_{t=0}^{\infty} \beta^t E_{i_0} (2\lambda\theta\delta_t - 2\lambda\theta^2(s_t^2 + \delta_t^2)) - \frac{1}{2}\tilde{\gamma}\sigma_\varepsilon^2 i_t^2 \quad (16)$$

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

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:

$$v_i = \max_{\{s, \delta\}} E_{i_0} (2\lambda\theta\delta - 2\lambda\theta^2(s^2 + \delta^2) - \frac{1}{2}\tilde{\gamma}\sigma_\varepsilon^2 i^2 + \beta v_{i'}) \quad (18)$$

$$i' = i + 2\lambda\theta^2 s + \varepsilon. \quad (19)$$

Standard techniques yield the following solution:

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

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

$$\delta^* = \frac{1}{2\theta} \quad (22)$$

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

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

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 a first-order autoregressive process:

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

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 therefore constant across inventory states.¹⁷ This result critically depends on the assumption that public buy and sell volume are linear in prices (see equation (9) and (10)).

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 a model-implied competitive spread, and analyzes Pareto efficiency.

The model’s primitive parameters. Table 4 summarizes these results with Panel A repeating the empirical results in Table 3 needed for identification of the model’s primitive parameters. To separately identify the λ and θ parameters, which characterize investors’ demand for liquidity, trading volume and spread are needed. 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 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. Panels B and C of Table 4 report median estimates for the overall sample and by quintile. Estimates by quintile and year are available in the supplementary materials. Panels B and C of Table 4 report interquartile ranges ($Q_{0.75} - Q_{0.25}$) as the model-implied variables are nonlinear functions of the empirical estimates (with estimates in denominators) which creates numerical issues. Medians and interquartile ranges are robust to these issues.

¹⁷Zabel (1981) and Mildestein and Schleef (1983) find the spread is independent of inventory. Ho and Stoll (1981) find inventory has a very small effect on the spread.

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

[insert Table 4]

The private value rate λ is highest for the largest-cap stocks: \$58,910 per stock per day. It decreases monotonically to \$1,450 for the smallest-cap stocks. Private value dispersion $\frac{1}{\theta}$ is 174 basis points for the largest-cap stocks, 133 for the second largest-cap stocks, and increases monotonically to 307 basis points for the smallest-cap stocks. These private value dispersions are somewhat smaller than the 5% standard deviation that [Hollifield, Miller, Sandås, and Slive \(2006\)](#) report based on structural estimation of a limit order book model.

We calculate the coefficient of relative risk aversion (γ) as the absolute risk aversion coefficient ($\hat{\gamma}$) divided by the intermediary's net spread, which equals her average consumption labeled \bar{c} in (11). The median relative risk aversion declines monotonically from 0.22 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). Finally, the risk aversion might be underestimated as the analysis assumes the intermediary cannot hedge any of the non-systematic risk on her position. In practice hedging may be possible, for example, by taking an offsetting position in stocks in the same industry.

2.4 Model-implied reservation spread

Section 2.2 shows that the intermediary's optimal control policy decomposes into two orthogonal decisions: one on the size of the bid-ask spread and the other on the amount of price pressure to apply. This naturally suggests a reservation spread which is the narrowest spread satisfying the intermediary's participation constraint.¹⁹ The reservation spread can be calculated using the previous section estimates of the model's

¹⁹The reservation spread could be competitively implemented by assigning the right to be the intermediary based on a second-

primitive parameters which are estimated only from the price pressure dynamics and not the spread observed in the data. The reservation spread is useful to check whether the parameter estimates are reasonable by comparing the reservation spread and the observed spread and to decompose the trading costs of liquidity demanders.

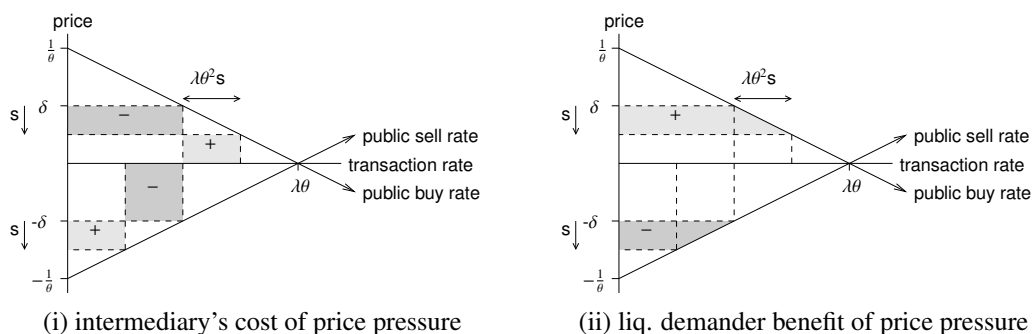
The annuity value of the inventory risk compensation equals $\beta P \sigma_\varepsilon^2$ if the intermediary starts with zero inventory (which is assumed for the remainder of the section). The model-implied conditional price pressure and inventory mean-reversion as described by equations (21) and (24) identify the factor βP . The remaining factor σ_ε^2 follows directly from inventory variance and mean-reversion, i.e., $\sigma_\varepsilon^2 = (1 - \rho_I^2) \sigma^2(I)$, where ρ_I is the first-order autocorrelation of inventory. This daily value divided by the intermediated daily volume naturally defines the ‘net spread’. The implied reservation spread is calculated by adding back the annuity value of the price pressure subsidy that the intermediary pays to mean-revert inventory. This value is calculated in Appendix III.

2.5 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)).

The right-hand side graph shows that liquidity demanders benefit since 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 this subsidy is obtained recursively (see Appendix III) and, when divided by daily intermediated volume, equals 0.56 basis points for the largest-cap stocks (7% of

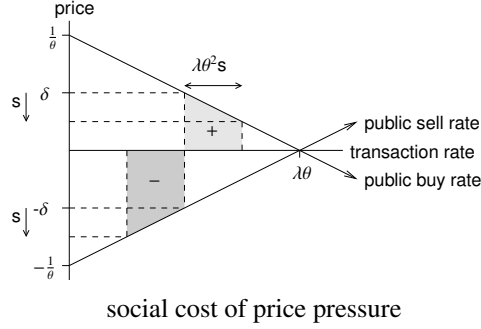
price auction among candidate intermediaries who bid based on a bid-ask spread commitment. The lowest-spread bidder is assigned the specialist role. In this case, all bid their reservation value which equals the spread that exactly compensates them for the utility cost of stochastic dividend on nonzero inventory positions, i.e., the second term in equation (20). In practice the NYSE did not assign intermediaries through such a procedure.



their reservation spread) and 15.13 basis points (17% of their reservation spread) for the smallest-cap stocks (see Panel C of Table 4).

The subsidy is paid by the intermediary so the gross spread is the sum of this subsidy and net spread. This is what the model implies one should observe in the data, i.e., this is the model-implied effective (half) spread. Panel C of Table 4 finds that it is 7.63 basis points for the largest-cap stocks and it increases monotonically to 87.41 basis points for the smallest-cap stocks. It is reassuring that the model-implied spread is of the same magnitude as the observed spread. It is slightly below for the largest-cap stocks (7.63 vs. 8.41 basis points) and increases to roughly 50% higher for the smallest-cap stocks (88.79 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-cap stocks where adverse selection is larger ([Hasbrouck \(1991\)](#)). The observed spread is net of the value of such privilege and, therefore, can be lower than the reservation spread.

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. 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 $\lambda\theta^2s^2$ as depicted in the following graph. Solving the ‘social planner dynamic program’ yields a solution similar to the intermediary’s solution where P in



equation (23) is replaced by \tilde{P} :

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

Panel C of Table 4 shows that the social planner’s implied ‘net spread’ is 6.20 basis points for the largest-cap stocks and increases to 54.48 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 0.61 basis points (10% of the net spread) for the largest-cap stocks and 17.66 basis points (32% of the net spread) for the smallest-cap stocks.

3 Discussion of 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 an advantageous central position in the trading process and imposed obligations upon them. This special role could affect the interpretation of inventory’s role in the price process. We first discuss the informational advantages and then other facets of the NYSE trading structure.

3.1 NYSE informational structure

An important privilege of the specialist at the time of our sample is the last-mover advantage on incoming orders, i.e., the specialist could decide on any incoming order to either trade against it herself by offering price improvement or to route it to the limit order book. Most importantly, in making this decision the

specialist could condition on some information about the identity of the order submitter, e.g., the brokerage firm handling the order. Her rationally exploiting this privilege by trading against uninformed orders and sending the informed orders to the book is often referred to as cream skimming (see, e.g., Benveniste, Marcus, and Wilhelm (1992), Rock (1990), Seppi (1997), and Parlour and Seppi (2003)). Cream skimming should reduce the adverse selection faced by the specialist which is an important reason for why this paper focuses on inventory control and price effects.

In the presence of cream skimming, if market participants can observe who is trading with whom, trades that involve the specialist should have smaller immediate price impacts, reflecting their lesser information content. A specialist would take on an inventory position and there would then be an immediate price reversal. But if other market participants cannot observe when the specialist trades, all trades would have similar immediate price impacts. The price reversals would not be immediate and would occur only as the underlying information (or lack thereof) is revealed.

While inventory control and cream skimming can both predict that inventories negatively correlate with the transitory component of prices, they make other differing predictions. First, unlike inventory control, cream skimming does not predict that prices have an unconditional transitory component. The pricing errors observed in the data through, for example, the negative autocorrelations in daily midquote returns (see Table 2) is predicted by inventory control and not by cream skimming. Second, the empirical evidence shows that the specialist inventories are negatively correlated with the informational (random-walk) component of price changes for all quintiles and significantly so except for the smallest-cap quintile (see Panels F and G in Table 2). In the state space model, this effect shows up through the negative coefficient on surprise inventory in the efficient price equation (see Panel C in Table 3). The adverse selection of the specialist suggests that cream skimming at the daily horizon is imperfect at best.

Finally, inventory control predicts that intermediaries pay a subsidy to reduce inventory risk. Hence, the information and risk management explanations make opposite predictions for the profitability of low-frequency (daily) position changes: cream skimming predicts positive profits on daily position changes while inventory control predicts negative profits. In a period before our sample Hasbrouck and Sofianos (1993) find that NYSE-specialist gross profits are positive on short- and medium-term positions and negative on long-term positions. Studying how the 2001 change to decimalization affected specialist gross profits Coughenour and Harris (2004) find negative specialist gross profits for medium and large stocks at horizons

greater than half a day. Hence, the analysis of specialist gross profits is consistent with inventory control and not cream skimming.

3.2 Non-informational NYSE market structure

In exchange for the informational advantages discussed above the NYSE imposed obligations upon 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 accommodate investor liquidity demands via taking inventory positions to prevent “transaction prices from overshooting beyond their equilibrium levels.” If this prevention of overshooting leads to prices adjusting to their efficient prices more slowly, then it causes inventories to be positively associated with pricing errors which is 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 likely that other investors compete with the specialist by placing limit orders to supply liquidity. This is especially important in the NYSE’s 2007 market structure changes (after our sample period) which resulted in a reduced role for the specialist (Hendershott and Moulton (2009)). Additional liquidity suppliers should reduce net spread and price pressure by increasing the risk bearing capacity in the intermediation sector. The most efficient manner to share risk among intermediaries 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 would

be adjusted by the reciprocal of the adjustment to the conditional price pressure.

4 Conclusion

Empirically we use twelve years of NYSE intermediary data to estimate time series price dispersion due to intermediaries' inventories in financial markets. These price pressures are deviations of prices from fundamental values due to the inventory risks born by intermediaries providing liquidity to facilitate continuous marketing clearing for asynchronously arriving investors. We develop an empirical model to separate the impact of intermediaries' inventory from price effects due to information or permanent shocks to supply and demand. 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 the distribution of investors' private values. We find:

1. A \$100,000 inventory shock causes price pressure of 0.98% 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 is large: 1.18% and 0.17% for small and large stocks, respectively. For small stocks the ratio of transitory volatility to the permanent volatility is greater than one, implying that daily price changes are driven more by 'noise' from price pressures than from information about fundamentals.
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 3.23 basis points of the value traded, which is approximately 10% of the cost of immediacy.

The significant size of price pressure suggests 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.

While these financial market price pressures are substantial, especially in small stocks, the effects dissipate within days or weeks. In decentralized markets for other goods and commodities production and consumption decisions likely adjust more slowly and search costs are higher. These both should make intermediaries' inventories have larger and longer lasting price effects.

Appendix I: Details on the likelihood optimization

The likelihood of the state space model described by equations (3), (4), and (5) is optimized in essentially three steps so as to minimize the probability of finding a local maximum. The optimization is implemented in **Ox** using standard optimization routines. It uses the Kalman filter and smoother from **ssfpack** which is an add-on package in **Ox** (see [Koopman, Shephard, and Doornik \(1999\)](#)).

1. An OLS regression of log price difference on contemporaneous and lagged f_t yields starting values for β_i and $\beta_i^0, \dots, \beta_i^k$ (see equations (3) and (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 β s and $(\sigma(\varepsilon), \varphi)$. The latter runs over a nine by nine grid where 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 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 its upper bound of and price pressure variance approaches the stock's unconditional variance. The optimizer starts to load the observed price series on two nonstationary series, i.e., the efficient price and the price pressure, and becomes unstable.

The Kalman filter is initialized with a diffuse distribution for the unobserved efficient price m_0 and the unconditional price pressure distribution for s_0 , i.e., $s_0 \sim N(0, \frac{\sigma^2(\varepsilon)}{1-\varphi^2})$.

3. The likelihood is optimized where all parameters are free and starting values for $(\beta_i, \beta_i^0, \dots, \beta_i^k, \sigma(\varepsilon), \varphi)$ 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 with respect to its 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 account for the fact that not all investor trading occurs via the intermediary. The intermediated volume, denoted ν in the equations below, is set equal to the fraction of trading volume in which the specialist is a participant: *specialist_particip* times total trading volume *dollar_volume*. The effective (half) spread is denoted δ^e . The expressions for the model's primitive parameters are:

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

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

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

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

The spread analysis uses:

$$net_spread = \frac{-\alpha * (1 - \rho_I^2) * \sigma^2(I)}{\rho_I * \nu}, \quad (29)$$

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

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

where

$$RQ \equiv \left(\frac{1}{\beta * \rho_I} - 1\right) * (1 - \rho_I). \quad (32)$$

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 off on an inventory of i . The Markovian law of motion for the intermediary's inventory position allows for

Bellmanizing this value as follows:

$$w_i = E_i[\frac{1}{2}Qs^2 + \beta w_i]. \quad (33)$$

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

$$A + Bi^2 = w_i = \frac{1}{2}Q\alpha^2 i^2 + \beta E_i[A + B(i + Q\alpha i + \varepsilon)^2] = (\beta A + \beta B\sigma_\varepsilon^2) + (\frac{1}{2}Q\alpha^2 + \beta B(1 + Q\alpha)^2)i^2. \quad (34)$$

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

References

- Barsky, R., T. Juster, M. Kimball, and M. Shapiro (1997). Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement survey. *Quarterly Journal of Economics* 112, 537–579.
- Benveniste, L. M., A. J. Marcus, and W. J. Wilhelm (1992). What’s special about the specialist? *Journal of Financial Economics* 32, 61–86.
- Brunnermeier, M. K. and L. H. Pedersen (2009). Market liquidity and funding liquidity. *Review of Financial Studies* 22, 2201–2238.
- 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.
- Coughenour, J. F. and L. E. Harris (2004). Specialist profits and the minimum price increment. Manuscript, University of Southern California.
- 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. (2010). Presidential address: Asset price dynamics with slow-moving capital. *Journal of Finance* 65, 1237–1267.
- Duffie, D., N. Gârleanu, and L. Pedersen (2005). Over-the-counter markets. *Econometrica* 73, 1815–1847.
- 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.
- 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 (2009). Automation, speed, and stock market quality: The nyse’s 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.
- Hollifield, B., R. A. Miller, P. Sandås, and J. Slive (2006). Estimating the gains from trade in limit order markets. *Journal of Finance* 61, 2753–2804.
- 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.
- Parlour, C. and D. Seppi (2003). Liquidity-based competition for order flow. *Review of Financial Studies* 16(2), 301–343.
- Pastor, L. and R. Stambaugh (2003). Liquidity risk and expected returns. *Journal of Political Economy* 111(3), 642–685.
- Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22, 435–480.
- 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.
- Seppi, D. (1997). Liquidity provision with limit orders and a strategic specialist. *Review of Financial Studies* 10, 103–150.
- 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 that merges data from 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 the average market capitalization where quintile 1 contains the largest-cap stocks (i.e., a stock is in the same quintile for all years in the sample).

variable	description (units)	source	mean Q1	mean Q2	mean Q3	mean Q4	mean Q5	st. dev.	wi-thin ^a
$midquote_{it}$	closing midquote, div/split adjusted ^b (\$)	NYSE	53.76	44.54	36.65	28.58	19.21		22.21
$invent_share_{it}$	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	412.61	168.98	129.48	75.44	77.90		1,383.32
$shares_outst_{it}$	shares outstanding (million)	CRSP	729.94	157.73	70.08	36.27	18.73		283.85
$market_cap_{it}$	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., $x_{it}^* = x_{it} - \bar{x}_i$.

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

Table 2: Descriptive statistics midquote price and inventory dynamics

This table presents descriptive statistics on daily log midquote price change and intermediary inventory. It is meant to illustrate these series' dynamics and their interaction building up to the state space model of Table 3. A common factor has been removed from both series: for price changes the series is regressed on market return and the residuals enter the analysis; for inventory, the series is regressed on a common factor that is defined as the cross-sectional mean each day of the standardized inventory series. In Panels F and G, the series' innovations are identified as the residuals from an AR(5) model. The table reports average estimates for each stock each year by market-cap quintile and year. The t -statistic pertains to the 'all' (column) parameter estimates and is based on a robust standard error obtained by stock and year clustering (see, e.g., [Petersen \(2009\)](#)).

	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	all	t -stat
<i>Panel A: First order autocorrelation idiosyncratic log price change</i>														
Q1	-0.03	-0.03	-0.04	-0.08	-0.02	-0.01	0.01	-0.00	-0.05	-0.01	-0.00	-0.02	-0.02	-3.35
Q2	-0.02	-0.02	-0.04	-0.06	-0.01	-0.02	-0.01	0.00	-0.05	-0.02	-0.02	-0.03	-0.02	-4.32
Q3	-0.00	-0.01	-0.03	-0.06	-0.02	-0.04	-0.03	-0.01	-0.06	-0.05	-0.03	-0.04	-0.03	-5.80
Q4	-0.02	-0.02	-0.03	-0.06	-0.01	-0.03	-0.05	-0.03	-0.08	-0.06	-0.08	-0.05	-0.04	-6.19
Q5	-0.05	-0.05	-0.02	-0.00	0.03	0.01	-0.02	0.01	-0.01	-0.02	-0.04	-0.02	-0.02	-2.03
all	-0.02	-0.03	-0.03	-0.05	-0.01	-0.02	-0.02	-0.01	-0.05	-0.03	-0.03	-0.03	-0.03	-6.66
<i>Panel B: Second order autocorrelation idiosyncratic log price change</i>														
Q1	-0.03	-0.05	-0.04	-0.02	-0.02	-0.02	-0.06	-0.05	-0.01	-0.01	-0.02	-0.01	-0.03	-6.35
Q2	-0.03	-0.03	-0.03	-0.01	-0.03	-0.01	-0.02	-0.02	-0.00	-0.02	-0.01	-0.02	-0.02	-5.91
Q3	-0.02	-0.02	-0.02	-0.02	-0.01	-0.00	-0.01	-0.02	0.01	-0.00	-0.02	-0.01	-0.01	-4.71
Q4	-0.01	-0.02	-0.02	-0.03	-0.02	-0.00	-0.01	-0.01	0.01	-0.01	0.01	-0.00	-0.01	-3.20
Q5	-0.01	-0.01	-0.01	-0.01	-0.00	0.00	-0.01	0.00	-0.00	-0.01	0.00	0.00	-0.00	-2.43
all	-0.02	-0.02	-0.02	-0.01	-0.02	-0.01	-0.02	-0.02	-0.00	-0.01	-0.01	-0.01	-0.01	-6.36
<i>Panel C: Standard deviation of idiosyncratic component specialist inventory I_{it} (in \$1000)</i>														
Q1	691	968	813	964	1126	1336	1344	1489	1472	1122	1118	1128	1131	11.06
Q2	473	510	489	524	530	695	820	647	448	391	441	400	531	12.26
Q3	374	429	383	372	430	452	668	437	294	242	266	271	385	10.65
Q4	226	254	255	261	291	315	320	333	229	163	145	147	245	11.95
Q5	167	159	168	235	205	224	211	187	130	109	95	96	166	11.31
all	386	464	421	471	516	605	672	619	514	406	414	408	491	14.68
<i>Panel D: First order autocorrelation idiosyncratic component specialist inventory I_{it}</i>														
Q1	0.28	0.27	0.26	0.22	0.25	0.27	0.28	0.29	0.28	0.34	0.37	0.25	0.28	21.24
Q2	0.47	0.46	0.44	0.38	0.35	0.34	0.36	0.32	0.25	0.28	0.33	0.25	0.35	15.97
Q3	0.59	0.59	0.56	0.51	0.49	0.45	0.41	0.41	0.30	0.31	0.34	0.24	0.43	12.68
Q4	0.74	0.73	0.71	0.66	0.63	0.63	0.59	0.57	0.40	0.38	0.36	0.29	0.56	12.56
Q5	0.82	0.80	0.80	0.77	0.78	0.79	0.77	0.76	0.66	0.61	0.57	0.51	0.72	23.89
all	0.58	0.57	0.56	0.51	0.50	0.50	0.48	0.47	0.38	0.38	0.39	0.31	0.47	19.06
<i>Panel E: Regression coef of log price change on lagged inventory, both idiosyncratic (in bps per \$1000)</i>														
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	8.71
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	10.05
Q3	0.02	0.02	0.02	0.02	0.02	0.03	0.04	0.01	0.05	0.03	0.02	0.03	0.03	7.49
Q4	0.03	0.02	0.04	0.05	0.04	0.06	0.07	0.04	0.11	0.06	0.07	0.06	0.05	7.43
Q5	0.08	0.07	0.03	0.03	0.07	0.06	0.08	0.08	0.15	0.09	0.10	0.13	0.08	6.63
all	0.03	0.03	0.02	0.02	0.03	0.03	0.05	0.03	0.07	0.04	0.04	0.05	0.04	9.30
<i>Panel F: Regression coef of log price innov on lagged inventory, both idiosyncratic (in bps per \$1000)</i>														
Q1	-0.03	-0.03	-0.04	-0.07	-0.01	0.00	0.02	-0.00	-0.05	-0.00	0.00	-0.01	-0.02	-2.60
Q2	-0.02	-0.02	-0.03	-0.07	-0.01	-0.00	0.00	0.02	-0.05	-0.02	-0.02	-0.03	-0.02	-3.10
Q3	-0.00	-0.00	-0.03	-0.06	-0.02	-0.02	-0.02	-0.01	-0.05	-0.05	-0.03	-0.04	-0.03	-4.95
Q4	-0.02	-0.03	-0.03	-0.06	-0.00	-0.02	-0.04	-0.03	-0.08	-0.07	-0.08	-0.04	-0.04	-5.78
Q5	-0.06	-0.05	-0.03	-0.00	0.04	0.02	-0.03	0.02	-0.01	-0.02	-0.04	-0.02	-0.02	-1.76
all	-0.03	-0.03	-0.03	-0.05	-0.00	-0.01	-0.01	-0.00	-0.05	-0.03	-0.03	-0.03	-0.02	-5.22
<i>Panel G: Regression coef of log price innov on lagged inventory innov, both idiosyncratic (in bps per \$1000)</i>														
Q1	-0.03	-0.03	-0.04	-0.08	-0.02	-0.00	0.02	-0.01	-0.05	-0.00	-0.00	-0.01	-0.02	-2.84
Q2	-0.02	-0.02	-0.03	-0.07	-0.01	-0.01	0.00	0.01	-0.05	-0.02	-0.02	-0.03	-0.02	-3.40
Q3	-0.00	0.00	-0.03	-0.06	-0.02	-0.02	-0.02	-0.02	-0.06	-0.04	-0.03	-0.04	-0.03	-4.81
Q4	-0.02	-0.03	-0.03	-0.06	0.00	-0.02	-0.04	-0.04	-0.08	-0.07	-0.08	-0.04	-0.04	-5.33
Q5	-0.06	-0.04	-0.01	-0.01	0.04	0.01	-0.03	0.03	-0.01	-0.03	-0.03	-0.02	-0.01	-1.64
all	-0.03	-0.02	-0.03	-0.05	-0.00	-0.01	-0.01	-0.00	-0.05	-0.03	-0.03	-0.03	-0.03	-5.30

Table 3: State space model estimates

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

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

where i indexes stocks and t indexes days, m_{it} is the end-of-day unobserved efficient price, s_{it} is the pricing error, δ_t is the required return obtained as the (time-varying) riskfree rate plus beta times a 6% risk premium, f_t is the demeaned market return, p_{it} is the end-of-day observed midquote, I_{it} is the idiosyncratic component of the specialist end-of-day inventory position, \hat{I}_{it} is the ('surprise') innovation in this position as obtained from an AR(9) model, β_i^j captures potential lagged adjustment to market returns (known to exist for small-cap stocks), and w_{it} and ε_{it} are mutually independent i.i.d. error terms. The model is estimated using maximum likelihood where the error terms w_{it} and ε_{it} are assumed to be normally distributed. The table reports average estimates for each stock each year by market-cap quintile and year. The t -statistic pertains to the 'all' (column) parameter estimates and is based on a robust standard error obtained by stock and year clustering (see, e.g., Petersen (2009)).

	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	all	t -stat
<i>Panel A: α_i conditional price pressure (in bps per \$1000)</i>														
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	-11.03
Q2	-0.06	-0.05	-0.05	-0.04	-0.03	-0.04	-0.04	-0.05	-0.04	-0.03	-0.03	-0.03	-0.04	-8.92
Q3	-0.12	-0.10	-0.09	-0.07	-0.09	-0.11	-0.10	-0.06	-0.09	-0.06	-0.06	-0.06	-0.08	-10.78
Q4	-0.33	-0.26	-0.29	-0.27	-0.26	-0.29	-0.25	-0.23	-0.26	-0.14	-0.18	-0.15	-0.24	-10.49
Q5	-1.05	-1.04	-0.76	-0.76	-1.02	-0.98	-1.06	-1.24	-1.19	-0.86	-0.93	-0.87	-0.98	-9.95
all	-0.32	-0.29	-0.24	-0.23	-0.28	-0.29	-0.30	-0.32	-0.32	-0.22	-0.24	-0.23	-0.27	-11.01
<i>Panel B: $\alpha_i \sigma(I)_i$ explained transitory volatility, i.e., unconditional price pressure (in bps)</i>														
Q1	14	13	13	16	19	20	25	30	22	15	14	8	17	
Q2	26	24	21	19	20	30	35	31	24	14	15	13	23	
Q3	39	40	35	31	39	45	44	35	23	16	19	17	32	
Q4	65	64	68	57	63	75	75	65	40	21	23	20	53	
Q5	112	108	113	107	151	170	179	154	107	92	70	54	118	
all	51	50	50	46	58	68	72	63	43	31	28	22	49	
<i>Panel C: κ_i informativeness order imbalance innovation \hat{I} (in bps per \$1000)</i>														
Q1	-0.08	-0.07	-0.07	-0.07	-0.08	-0.08	-0.08	-0.07	-0.07	-0.06	-0.05	-0.05	-0.07	-18.01
Q2	-0.16	-0.13	-0.13	-0.14	-0.17	-0.16	-0.17	-0.14	-0.16	-0.15	-0.13	-0.13	-0.15	-20.01
Q3	-0.22	-0.18	-0.17	-0.21	-0.25	-0.24	-0.24	-0.24	-0.25	-0.25	-0.25	-0.22	-0.23	-24.22
Q4	-0.34	-0.27	-0.29	-0.36	-0.43	-0.40	-0.45	-0.42	-0.52	-0.53	-0.49	-0.48	-0.42	-13.81
Q5	-0.58	-0.42	-0.61	-0.61	-0.69	-0.58	-0.76	-0.87	-1.20	-1.37	-1.21	-1.33	-0.85	-8.63
all	-0.27	-0.21	-0.26	-0.28	-0.32	-0.29	-0.34	-0.35	-0.44	-0.47	-0.42	-0.44	-0.34	-12.82
<i>Panel D: $\kappa_i \sigma(\hat{I})_i$ explained permanent volatility (in bps)</i>														
Q1	38	36	38	44	57	69	75	55	55	36	29	26	47	
Q2	41	38	40	45	64	75	83	61	56	43	35	34	51	
Q3	43	40	41	54	70	74	79	74	58	47	40	34	54	
Q4	47	39	42	52	70	70	81	73	65	57	46	41	57	
Q5	43	38	49	55	67	65	77	90	79	74	62	59	63	
all	42	38	42	50	66	71	79	70	63	52	42	39	54	
<i>Panel E: $\sigma(w)_i$ permanent volatility (in bps)</i>														
Q1	124	124	129	135	176	202	256	194	188	127	107	103	156	
Q2	141	135	138	143	185	217	269	213	206	150	125	123	170	
Q3	149	141	145	156	201	216	257	212	204	151	136	133	175	
Q4	161	151	153	156	198	215	242	214	199	161	143	148	178	
Q5	174	168	179	179	225	235	263	247	243	200	175	186	206	
all	150	144	149	154	197	217	257	216	208	158	137	139	177	

<continued on next page>

<continued from previous page>

	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	all	t-stat
<i>Panel F: $\frac{\alpha_i^2 \sigma^2(I)_i}{\sigma^2(w)_i + \beta_i^2 \sigma^2(\bar{f})}$ ratio of price pressure variance and permanent variance</i>														
Q1	0.02	0.02	0.02	0.01	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	0.03	0.03	0.03	0.02	0.01	0.02	0.01	0.03	
Q3	0.29	0.25	0.16	0.08	0.11	0.16	0.05	0.08	0.02	0.01	0.03	0.07	0.11	
Q4	0.31	0.47	0.48	0.26	0.22	0.37	0.24	0.18	0.05	0.03	0.03	0.05	0.22	
Q5	1.09	1.35	1.13	1.11	1.23	2.40	1.82	1.15	0.79	2.08	0.59	0.59	1.28	
all	0.36	0.43	0.37	0.30	0.32	0.59	0.43	0.29	0.18	0.43	0.14	0.15	0.33	
<i>Panel G: $\frac{\alpha_i^2 \sigma^2(I)_i}{\sigma^2(w)_i}$ ratio of price pressure variance and permanent 'idiosyncratic' variance</i>														
Q1	0.03	0.03	0.02	0.02	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	0.03	0.04	0.04	0.03	0.02	0.03	0.02	0.04	
Q3	0.35	0.28	0.20	0.11	0.17	0.18	0.05	0.10	0.04	0.02	0.05	0.11	0.14	
Q4	0.37	0.49	0.54	0.31	0.30	0.40	0.25	0.23	0.08	0.04	0.05	0.05	0.26	
Q5	1.25	1.46	1.25	1.28	1.55	2.49	2.04	1.43	1.00	2.32	0.72	0.65	1.45	
all	0.42	0.47	0.41	0.35	0.41	0.62	0.48	0.37	0.24	0.48	0.18	0.17	0.38	
<i>Panel H: $\sigma(\varepsilon)$; error term transitory volatility (in bps)</i>														
Q1	20	17	23	30	24	20	26	19	32	18	12	15	21	
Q2	20	20	22	28	24	22	34	18	39	23	17	19	24	
Q3	19	18	23	24	28	29	35	26	41	26	20	22	26	
Q4	29	26	24	27	18	23	37	29	42	30	30	25	28	
Q5	43	42	34	25	22	29	43	35	38	28	29	25	33	
all	26	24	25	27	23	24	35	25	39	25	21	21	26	

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

This table presents an economic analysis of 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 intermediary's coefficient of relative risk aversion (γ) which is the model's absolute risk aversion ($\bar{\gamma}$) divided by a competitive intermediary's average consumption. The conditional price pressure estimates along with other trading variables measured in the empirical part of the paper (Panel A contains sample averages taken from Tables 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
<i>Panel A: Measured variables that identify model's primitive parameters (cf. Table 1 and 3)</i>						
conditional price pressure α_i (in bps per \$1000)	-0.02	-0.04	-0.09	-0.25	-0.99	-0.28
stdev daily inventory $\sigma(I)$ (in \$1000)	1131	530	385	245	166	491
1st order autocorrelation inventory	0.28	0.35	0.43	0.56	0.72	0.47
price risk inventory $\sigma(w)_i$ (in bps)	153	170	176	179	207	177
daily dollar volume ^a (in \$million)	10.86	2.98	1.43	0.60	0.21	3.22
effective half spread (in bps)	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
<i>Panel B: Identification of model's primitive parameters</i>						
daily private value rate λ (in \$1000)	58.91	14.43	6.29	2.83	1.45	5.90
dispersion private value $\frac{1}{\theta}$ (in bps)	(236.50) 174	(36.51) 133	(12.80) 146	(4.76) 182	(2.28) 307	(21.91) 185
intermediary's relative risk aversion γ	(297) 0.22	(197) 0.13	(164) 0.10	(197) 0.08	(362) 0.05	(256) 0.10
	(1.01)	(0.54)	(0.34)	(0.25)	(0.13)	(0.34)
<i>Panel C: Decomposition of the spread paid by liquidity demander</i>						
(1) model-implied competitive spread (in bps)	7.63	14.32	21.52	42.30	87.41	25.18
	(17.20)	(30.03)	(44.83)	(81.57)	(178.75)	(66.49)
(2) price pressure subsidy to liq demander (in bps)	0.56	1.40	2.54	5.86	15.13	2.82
	(1.11)	(3.05)	(5.62)	(13.32)	(33.03)	(9.38)
(3) net spread to liq demander ^c (1)-(2) (in bps)	6.97	12.47	19.09	36.34	72.21	22.07
	(15.69)	(26.13)	(39.77)	(69.13)	(145.93)	(56.59)
(4) constrained Pareto efficient spread (in bps)	6.20	10.76	16.01	29.28	54.48	18.41
	(14.09)	(22.14)	(33.10)	(54.24)	(108.47)	(44.68)
(5) deadweight loss ^c (3)-(4) (in bps)	0.61	1.56	2.83	6.58	17.66	3.16
	(1.19)	(3.41)	(6.35)	(15.58)	(38.62)	(10.88)

^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

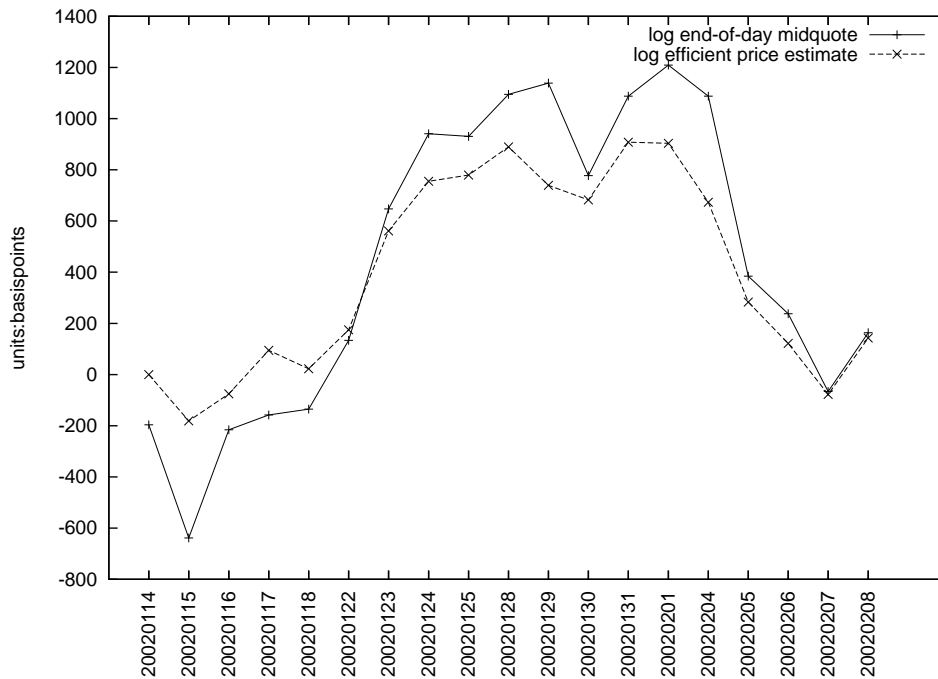
^c: the reported difference is the median of differences, not the difference of (reported) medians

Figure 1: Price change decomposition conditional on the state space model estimates, not using inventory data

These figures depict the observed end-of-day log (midquote) price and the unobserved efficient price estimate for a representative stock (Rex Stores Corporation, ticker RSC, CRSP PERMNO 68830). The efficient price estimate is calculated with the Kalman smoother (and thus takes all observations into account) after estimating the following state space model:

$$\begin{aligned}
 \text{(observed midquote price)} \quad p_{it} &= m_{it} + s_{it} \\
 \text{(unobserved efficient price)} \quad m_{it} &= m_{i,t-1} + \delta_{i,t} + \beta_i f_t + w_{it} \\
 \text{(unobserved transitory price deviation)} \quad s_{it} &= \varphi_i s_{i,t-1} + \beta_i^0 f_t + \dots + \beta_i^3 f_{t-3} + \varepsilon_{it}
 \end{aligned}$$

where i indexes stocks and t indexes days, m_{it} is the end-of-day unobserved efficient price, s_{it} is the pricing error, δ_t is the required return obtained as the (time-varying) riskfree rate plus beta times a 6% risk premium, f_t is the demeaned market return, p_{it} is the end-of-day observed midquote, β_i^j captures potential lagged adjustment to market returns (known to exist for small-cap stocks), and w_{it} and ε_{it} are mutually independent i.i.d. error terms. The model is estimated using maximum likelihood 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 pricing error — the difference between the observed price and the efficient price — against the idiosyncratic inventory position of the specialist. The third panel graphs the efficient price innovation against the contemporaneous ('surprise') idiosyncratic inventory innovation.



<continued on next page>

<continued from previous page>

