

Trading Costs and Returns for U.S. Equities: Estimating Effective Costs from Daily Data

Author(s): Joel Hasbrouck

Source: *The Journal of Finance*, Vol. 64, No. 3 (Jun., 2009), pp. 1445-1477

Published by: Wiley for the American Finance Association

Stable URL: <https://www.jstor.org/stable/20488006>

Accessed: 19-12-2018 03:07 UTC

## REFERENCES

Linked references are available on JSTOR for this article:

[https://www.jstor.org/stable/20488006?seq=1&cid=pdf-reference#references\\_tab\\_contents](https://www.jstor.org/stable/20488006?seq=1&cid=pdf-reference#references_tab_contents)

You may need to log in to JSTOR to access the linked references.

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



American Finance Association, Wiley are collaborating with JSTOR to digitize, preserve and extend access to *The Journal of Finance*

# Trading Costs and Returns for U.S. Equities: Estimating Effective Costs from Daily Data

JOEL HASBROUCK\*

## ABSTRACT

The effective cost of trading is usually estimated from transaction-level data. This study proposes a Gibbs estimate that is based on daily closing prices. In a validation sample, the daily Gibbs estimate achieves a correlation of 0.965 with the transaction-level estimate. When the Gibbs estimates are incorporated into asset pricing specifications over a long historical sample (1926 to 2006), the results suggest that effective cost (as a characteristic) is positively related to stock returns. The relation is strongest in January, but it appears to be distinct from size effects.

INVESTIGATIONS INTO THE ROLE of liquidity and transaction costs in asset pricing must generally confront the fact that while many asset pricing tests make use of U.S. equity returns from 1926 onward, the high-frequency data used to estimate trading costs are usually not available prior to 1983. Accordingly, most studies either limit the sample to the post-1983 period of common coverage or use the longer historical sample with liquidity proxies estimated from daily data. This paper falls into the latter group. Specifically, I propose a new approach to estimating the effective cost of trading and the common variation in this cost. These estimates are then used in conventional asset pricing specifications with a view to ascertaining the role of trading costs as a characteristic in explaining expected returns.<sup>1</sup>

\*Hasbrouck is with the Stern School of Business, New York University. For comments and suggestions I am grateful to the editor, the referee, Yakov Amihud, Ľuboš Pástor, Bill Schwert, Jay Shanken, Kumar Venkataraman, Sunil Wahal, and seminar participants at the University of Rochester, the NBER Microstructure Research Group, the Federal Reserve Bank of New York, Yale University, the University of Maryland, the University of Utah, Emory University, and Southern Methodist University. All errors are my own responsibility. Earlier versions of this paper and an SAS data set containing the long-run Gibbs sampler estimates are available on my web site at [www.stern.nyu.edu/~jhasbrou](http://www.stern.nyu.edu/~jhasbrou).

<sup>1</sup>Recent asset pricing analyses based on samples in which high-frequency data are available include Brennan and Subrahmanyam (1996), Easley, Hvidkjaer, and O'Hara (2002), Sadka (2004), and Korajczyk and Sadka (2008). Analyses that use proxies based on daily data include Amihud (2002), Pástor and Stambaugh (2003), Acharya and Pedersen (2005), and Spiegel and Wang (2005). Closing daily or annual bid-ask quotes are sometimes available over samples longer than those of the high-frequency data. Studies that use closing spreads include Stoll and Whalley (1983), Amihud and Mendelson (1986), Eleswarapu and Reinganum (1993), Reinganum (1990), and Chalmers and Kadlec (1998). Easley and O'Hara (2002) provide a broad survey of the links between asset pricing and microstructure.

For a buy order executed in a single trade, the effective cost is the execution price less the midpoint of the prevailing bid and ask quotes (and vice versa for a sale). In the simplest setting, the buyer is demanding immediacy by meeting the ask price set by a dealer or other liquidity supplier, and the effective cost represents the price of this immediacy. Admittedly, for more complicated strategies, particularly when the original order is executed over time in multiple trades, the effective cost does not generally account for the impact of earlier trades on subsequent prices. On the other hand, this measure is simple to compute (from detailed trade and quote records), easy to interpret, and widely used as an indicator of market quality.<sup>2</sup> Since 2000 the SEC has required U.S. market centers to report summary statistics of their effective costs based on the orders they actually receive and execute (Reg NMS Rule 605, formerly designated Rule 11ac1-5).

To estimate the effective cost from daily closing prices, I review the Roll (1984) model of price dynamics. Hasbrouck (2004) suggests a Bayesian Gibbs approach to the Roll model, and applies it to futures transaction data. This study generalizes the Hasbrouck model and applies it to daily CRSP U.S. equity data. The CRSP/Gibbs estimates are then compared to estimates based on high-frequency trade and quote (TAQ) data. This comparison sample spans from 1993 to 2005, and comprises roughly 300 firms per year (approximately 3,900 firm-years). In the comparison sample, the CRSP/Gibbs estimate of average effective cost achieves a correlation of 0.965 with the TAQ value. Overall, subject to some qualifications discussed in the body of the paper, these findings suggest that the CRSP/Gibbs estimates are strong proxies for the high-frequency measures. I therefore extend the estimates to the full CRSP daily sample (1926 to present).

Next I turn to applications of these proxies in asset pricing specifications. The earlier papers in this area view liquidity as a characteristic that drives a wedge between the returns an investor might realize net of trading costs and the gross returns used in most asset pricing tests (Amihud and Mendelson (1986)). This effect predicts a positive relation between gross returns and trading costs. Many of the studies cited in footnote 1 find such a link, but the evidence is mixed.

I find that, indeed, in diverse samples across listing venues and time, effective cost and returns are positively related. I also find, however, two problematic aspects to this relation. First, it is concentrated in January. This confirms, using a larger and broader sample, the seasonality results reported by Eleswarapu and Reinganum (1993). The effective cost seasonality appears to dominate the January small-firm effect. There is no obvious explanation, however, for this phenomenon. The second problem is that the coefficients on effective cost are too large to be consistent with the simplest trading stories. More precisely, the estimated impact of effective cost on returns can only be viewed as equilibrium compensation for trading expenses if the marginal agent's trading activity is substantially larger than the average measured over the sample period.

<sup>2</sup> Lee (1993) is representative of early work. Stoll (2006) is a recent example.

Beginning with Pástor and Stambaugh (2003), various studies examine the effects of dynamic liquidity variation and covariation. Pástor and Stambaugh find that liquidity risk, that is, the covariance between an asset's return and the common liquidity factor, is priced. Liquidity risk is also measured by Acharya and Pedersen (2005), using the Amihud (2002) illiquidity measure, and by Korajczyk and Sadka (2008), using high-frequency measures. This study implements a preliminary analysis of liquidity variation using the Gibbs estimates of effective cost. The results, however, are less supportive of liquidity risk effects.

The remainder of the paper is organized as follows. Section I describes the specification and computational procedures used to estimate effective cost. Data sources and sample construction are discussed in Section II. Section III examines the performance of the Gibbs estimates (relative to the TAQ values) in the comparison sample. Section IV discusses the long-run historical estimates of effective cost and trade directions for U.S. equities. Section V analyzes the link between returns and effective cost estimates in asset pricing specifications. Section VI discusses approaches to characterizing variation in effective costs, and implements an analysis of daily variation. Section VII concludes the paper.

## I. Bayesian Estimation of Effective Cost

### A. The Roll Model

Roll (1984) suggests a simple model of security prices in a market with transaction costs. The specifications estimated in this paper are variants of the Roll model, but the basic version is useful for describing the estimation procedure. The price dynamics may be stated as

$$\begin{aligned} m_t &= m_{t-1} + u_t \\ p_t &= m_t + cq_t, \end{aligned} \tag{1}$$

where  $m_t$  is the log quote midpoint prevailing prior to the  $t^{\text{th}}$  trade ("efficient price"),  $p_t$  is the log trade price, and the  $q_t$  are direction indicators, which take the values +1 (for a buy) or -1 (for a sale) with equal probability. The disturbance,  $u_t$ , reflects public information and is assumed to be uncorrelated with  $q_t$ . Roll motivates  $c$  as one-half the posted bid-ask spread, but since the model applies to transaction prices, it is natural to view  $c$  as the effective cost. The model has essentially the same form under time aggregation. In particular, although the model is sometimes estimated with transaction data (e.g., Schultz (2000)), it was originally applied to daily data, with  $q_t$  being interpreted as the direction variable for the last trade of the day.

The Roll model implies

$$\Delta p_t = m_t + cq_t - (m_{t-1} + cq_{t-1}) = c\Delta q_t + u_t, \tag{2}$$

from which it follows that  $c = \sqrt{-\text{Cov}(\Delta p_t, \Delta p_{t-1})}$ , where  $\text{Cov}(\Delta p_t, \Delta p_{t-1})$  is the first-order autocovariance of the price changes. The usual estimate of  $c$  is the sample analog of this quantity, and is called the "moment" or "autocovariance"

estimate because it uses a sample moment (the sample autocovariance) in lieu of the population value.

The moment estimate is feasible, however, only if the first-order sample autocovariance is negative. In samples of daily frequency this is often not the case. In annual samples of daily returns, Roll finds positive autocovariances in roughly half the cases. Harris (1990) discusses this and other aspects of this estimator. He shows that positive autocovariances are more likely for low values of the spread. Accordingly, one simple remedy is to assign an a priori value of zero. Another problem arises when there is no trade on a particular day, in which case CRSP reports the midpoint of the closing bid and ask. If these days are retained in the sample, the estimated cost will generally be biased downward, because the midpoint realizations do not include the cost. If these days are dropped from the sample, heteroskedasticity may arise since the efficient price innovations may span multiple days.

### B. Bayesian Estimation Using the Gibbs Sampler

Hasbrouck (2004) advocates a Bayesian approach. Completing the Bayesian specification requires specification of the distribution of  $u_t$ . I assume here that  $u_t \stackrel{d}{\sim} i.i.d. N(0, \sigma_u^2)$ . The prior distributions for parameters are discussed below.

The data sample is denoted  $p \equiv \{p_1, p_2, \dots, p_T\}$ . The unknowns comprise both the model parameters  $\{c, \sigma_u^2\}$  and the latent data, the trade direction indicators  $q \equiv \{q_1, \dots, q_T\}$ . (Knowing  $p$  and  $q$  suffices to determine  $m \equiv \{m_1, \dots, m_T\}$ .) The parameter posterior density  $f(c, \sigma_u^2 | p)$  is not obtained in closed form, but is instead characterized by random draws (from which means and other summary statistics may be computed). The random draws are generated using a Gibbs sampler whereby each unknown is drawn in turn from its full conditional (posterior) distribution. First,  $c$  and  $q$  are initialized to arbitrary feasible values. Next,  $c$  is drawn from  $f(c | \sigma_u^2, q, p)$ ;  $\sigma_u^2$  is drawn from  $f(\sigma_u^2 | c, q, p)$ ;  $q_1$  is drawn from  $f(q_1 | c, \sigma_u^2, q_2, q_3, \dots, q_T, p)$ , and so on.

The draws are described in more detail below, but one central feature of the model warrants emphasis. In the expression for  $\Delta p_t$  given by equation (2), if the  $\Delta q_t$  are known (or taken as given), the equation describes a simple linear regression with coefficient  $c$ . The linear regression perspective is a dominant theme of the present analysis. It simplifies implementation using standard results from Bayesian statistics, and suggests ways in which the model may be generalized.

#### B.1. Simulating the Coefficient(s) in a Linear Regression

The standard Bayesian normal regression model is  $y = Xb + e$  where  $y$  is a column vector of  $n$  observations of the dependent variable,  $X$  is an  $n \times k$  matrix of fixed regressors,  $b$  is a vector of coefficients, and the residuals are zero-mean multivariate normal  $e \sim N(0, \Omega_e)$ . Given  $\Omega_e$  and a normal prior on  $b$ ,  $b \sim N(\mu_b, \Omega_b)$ , the posterior is  $b \sim N(\mu_b^*, \Omega_b^*)$ , where  $\mu_b^* = (X'\Omega_e^{-1}X + \Omega_b^{-1})^{-1}(X'\Omega_e^{-1}y + \Omega_b^{-1}\mu_b)$  and  $\Omega_b^* = (X'\Omega_e^{-1}X + \Omega_b^{-1})^{-1}$ . Carlin and

Louis (2000), Lancaster (2004), and Geweke (2005) are contemporary textbook treatments.

In the present applications it is often necessary to impose inequality restrictions on the  $\beta$ . Typically, one or more coefficients is restricted to the positive domain. It is straightforward to show that when the  $b$  prior is restricted to  $\underline{b} < b < \bar{b}$ , the posterior has the same parameters as in the unrestricted case, but is truncated to the same interval as the prior (see, for example, Geweke (2005), Section 5.3.1). Hajivassiliou, McFadden, and Ruud (1996) discuss computationally efficient procedures for making random draws from truncated multivariate normal distributions.

### B.2. Simulating the Error Covariance Matrix

The primary results in this paper involve the case in which  $\Omega_e = \sigma^2 I$ . If the parameter prior is  $\sigma^2 \sim IG(\alpha, \beta)$ , where  $IG$  denotes the inverted gamma distribution, then the posterior is  $\sigma^2 \sim IG(\alpha^*, \beta^*)$ , where  $\alpha^* = \alpha + n/2$  and  $\beta^* = [\beta^{-1} + \sum e_i^2/2]^{-1}$ .

### B.3. Simulating the Trade Direction Indicators

The remaining step in the sampler involves drawing  $q \equiv \{q_1, \dots, q_T\}$  when  $c$  and  $\sigma_u^2$  are known. The procedure is sequential. The first draw is  $q_1 | q_2, \dots, q_T$ , the second draw is  $q_2 | q_1, q_3, \dots, q_T$ , the third draw is  $q_3 | q_1, q_2, q_4, \dots, q_T$ , etc., where the “|” notation denotes the conditional draw. The full set of conditioning information includes the price changes  $\Delta p \equiv \{\Delta p_2, \dots, \Delta p_T\}$  and the parameters  $c$  and  $\sigma_u^2$ .

The first realization of  $u_t$  to enter the observed prices is  $u_2$ . This may be written as a function of  $q_1$  according to  $u_2(q_1) = \Delta p_2 - cq_2 + cq_1$  (given  $q_2$ , etc.). A priori,  $u_2 \sim N(0, \sigma_u^2)$  and  $q_1 = \pm 1$  with equal probability. The posterior odds ratio of a buy versus a sell is

$$\frac{\Pr(q_1 = +1 | q_2, \dots)}{\Pr(q_1 = -1 | q_2, \dots)} = \frac{f(u_2(q_1 = +1))}{f(u_2(q_1 = -1))},$$

where  $f$  is the normal density function with mean zero and variance  $\sigma_u^2$ . The right-hand side of this is easily computed, and  $q_1$  is drawn using the implied (Bernoulli) probability.

To draw  $q_2$ , note that, given everything else, we may write  $u_2(q_2) = \Delta p_2 - cq_2 + cq_1$  and  $u_3(q_2) = \Delta p_3 - cq_3 + cq_2$ . Given the assumed serial independence of the  $u_t$ , the posterior odds ratio is

$$\frac{\Pr(q_2 = +1 | q_1, q_3, \dots)}{\Pr(q_2 = -1 | q_1, q_3, \dots)} = \frac{f(u_2(q_2 = +1))f(u_3(q_2 = +1))}{f(u_2(q_2 = -1))f(u_3(q_2 = -1))}.$$

Again, we compute the right-hand side and make the draw. In this fashion, we progress through the remaining  $q_t$ . For all draws of  $q_t$  (except the first and last) the posterior odds ratio involves only the adjacent disturbances  $u_t$  and  $u_{t+1}$ . The posterior odds ratio for the last draw is

$$\frac{\Pr(q_T = +1 | q_1, \dots, q_{T-1})}{\Pr(q_T = -1 | q_1, \dots, q_{T-1})} = \frac{f(u_T(q_T = +1))}{f(u_T(q_T = -1))}.$$

In some samples, for a subset of times, the trade directions may be known. These  $q_t$  may simply be left at their known values. A related situation arises from the CRSP convention of reporting quote midpoints on days with no trades. For these days we fix  $q_t = 0$ , implying that  $p_t = m_t$ , that is, that the quote midpoint is observed without error. This may be formally justified by positing a more general model that admits the possibility of no trade. If the no-trade probability is denoted  $\pi$ , for example, the general model would allow  $q_t$  to take on values 0, +1, and -1 with probabilities  $\pi$ ,  $(1 - \pi)/2$ , and  $(1 - \pi)/2$ , respectively. Assuming, however, that the no-trade days are known, that buys and sells are equally likely given a trade occurrence, and that we do not wish to estimate or characterize  $\pi$ , the more general model is observationally equivalent to the simpler one.

Another sort of observational equivalence is slightly more troublesome. It is natural to assume that trading costs are (at least on average) nonnegative, that is,  $c > 0$ . This is an economic assumption, however. From a statistical viewpoint, the model is observationally equivalent to one in which  $c < 0$  and all trade directions have the opposite signs ("buys" have  $q_t = -1$ , etc.). Simulated posteriors for  $c$  are therefore bimodal, symmetric about the origin. To rule out this "mirror" situation, it is convenient and sensible to impose the restriction  $c > 0$  on the prior.

Bayesian analyses sometimes use improper priors, often with the purpose of establishing an explicit connection to classical frequentist approaches. For example, letting  $\Omega_b^{-1}$  approach zero in the Bayesian regression model discussed above leads to posterior estimates that converge to the usual frequentist ones (e.g., Geweke, p. 81). The present situation does not admit this device, however. The regressors in equation (2) are the  $\Delta q_t$ , which are simulated. If the  $q_t$  drawn in one iteration (sweep) of the sampler all happen to have the same sign, then all of the  $\Delta q_t$  equal zero, and the computed regression is uninformative (for this sweep). In this case, a draw must be made from the prior distribution. Although this is an infrequent occurrence, it effectively rules out a prior for  $c$  that is proper but extremely diffuse.

### *C. The Basic Market-Factor Model and Sampler Specification*

The models estimated in this paper generalize on the basic Roll model in various respects. It is straightforward to add other regressors to equation (2). The motivation for doing so is that, intuitively, the procedure tries to allocate transaction price changes between "true" (efficient price) returns and transient trading costs. Anything that helps explain either component will sharpen the resolution. Return factors are obvious candidates for supplemental regressors. The basic market-factor model is

$$\Delta p_t = c \Delta q_t + \beta_m r_{mt} + u_t, \quad (3)$$

where  $r_{mt}$  is the market return on day  $t$ . It is assumed that the market return is independent of  $\Delta q_t$ . This would be the case if the trade direction indicators for the component securities are mutually independent, implying a diversification of bid-ask bounce. Note that although the present goal is improved estimation of  $c$ , it is likely that estimation of  $\beta_m$  will also be enhanced.

In the present applications (all involving U.S. equity data), the prior for  $c$  is the normal density with mean parameter equal to zero and variance parameter equal to  $0.05^2$  restricted to nonnegative values, denoted  $N^+(\mu = 0, \sigma^2 = 0.05^2)$ . The  $\mu$  and  $\sigma^2$  appearing here are only formal parameters: The actual mean and variance of the distribution will differ due to the truncation. The prior for  $\beta_m$  is  $N(\mu = 1, \sigma^2 = 1)$ ; that for  $\sigma_u^2$  is inverted gamma,  $IG(\alpha = 1 \times 10^{-12}, \beta = 1 \times 10^{-12})$ .

The sampler then follows the following program:

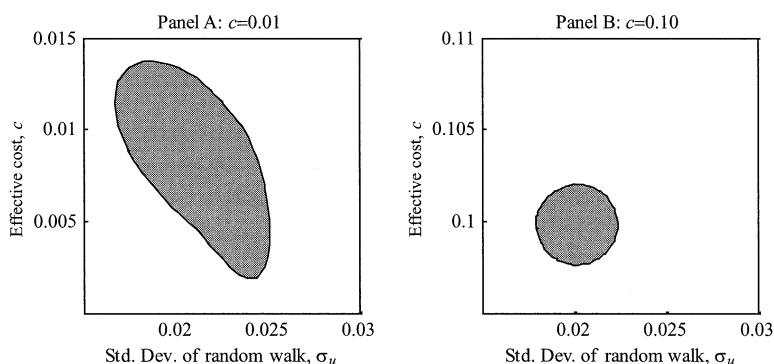
- Step 0 (initialization). Although the limiting behavior of the sampler is invariant to starting values, “reasonable” initial guesses may hasten convergence. The trade direction indicators  $q_t$  that do not correspond to midpoint reports are set to the sign of the most recent price change, with  $q_1$  set (arbitrarily) to  $+1$ ;  $\sigma_u^2$  is initially set to 0.0004 (roughly corresponding to a 30% annual idiosyncratic volatility). No initial values are required for  $c$  and  $\beta_m$ , as they are drawn first.
- Step 1. Based on the most recently simulated values for  $\sigma_u^2$  and the set of  $q_t$ , compute the posterior for the regression coefficients ( $c$  and  $\beta_m$ ) and make a new draw.
- Step 2. Given  $c$ ,  $\beta_m$ , and the set of  $q_t$ , compute the implied  $u_t$ , update the posterior for  $\sigma_u^2$ , and make a new draw.
- Step 3. Given  $c$ ,  $\beta_m$ , and  $\sigma_u^2$ , make draws for  $q_1, q_2, \dots, q_T$ . Go to Step 1.

To ease the computational burden, each sampler is run for only 1,000 sweeps. Although this value is small by the standards of most Markov chain Monte Carlo analyses, it appears to be sufficient in the present case, as experimentation with up to 10,000 sweeps does not materially affect the mean parameter estimates. Of the 1,000 draws for each parameter, the first 200 are discarded to “burn in” the sampler, that is, remove the effect of starting values. The average of the remaining 800 draws (an estimate of the posterior mean) is used as a point estimate of the parameter in subsequent analysis.

#### D. An Illustration

The essential properties of the estimator may be illustrated by considering two simulated price paths. The paths correspond to situations typical of U.S. equities. Both paths are of length 250 (roughly a year of daily observations). The standard deviation for the efficient price innovation is  $\sigma_u = 0.02$  (that is, about 2%, corresponding to an annual standard deviation of about 32%). For simplicity,  $\beta_m = 0$ . One simulated series of  $u_t$  and one simulated series of  $q_t$  are





**Figure 1. Posteriors for simulated price paths.** A quote-midpoint series of length 250 (roughly a year's worth of daily data) is simulated using volatility  $\sigma_u = 0.02$ ; 250 realizations are also generated for the trade direction indicators ( $q_t$ ). Using these values, two price series are simulated: one using an effective cost of  $c = 0.01$ , the other with  $c = 0.10$ . For each series, the joint parameter posterior is estimated using 10,000 draws of a Gibbs sampler. The shaded regions indicate the 90% confidence regions. In the two panels, the horizontal ( $\sigma_u$ ) axis and the scale of the vertical ( $c$ ) axis are identical.

used for both paths. The price paths are identical except for the scaling of the effective cost:  $c$  is either set to 0.01 or 0.10. The prior for  $c$  is  $N^+(0, 1)$ , that is, somewhat more diffuse than the prior used in the actual estimates. For each path the Gibbs sampler is run for 10,000 sweeps, with the first 2,000 discarded. The remaining 8,000 draws are used to characterize the posteriors.

Figure 1 illustrates the simulated 90% confidence regions for the parameter posteriors. Panel A (Panel B) depicts the posterior when  $c = 0.01$  ( $c = 0.10$ ). To facilitate comparisons, the horizontal axes ( $\sigma_u$ ) are identical. The vertical axes ( $c$ ) are shifted, but have the same scale.

The results are striking. In Panel A ( $c = 0.01$ ), the joint confidence region is large and negatively sloped. In Panel B ( $c = 0.10$ ), the confidence region is circular, centered around the population values, and compact.

To develop the intuition for this result, recall that the Gibbs procedure generates conditional random draws for the trade direction indicators. These draws characterize the posteriors for the trade direction indicators, and the sharpness of these posteriors corresponds very closely to what one might guess on the basis of looking at the price paths. When  $c$  is large relative to the efficient price increments, the price path appears distinctly “spikey” (with many reversals), as a consequence of the large bid-ask bounce. It is easy to confidently identify buys and sells, and the parameter posterior is concentrated. When  $c$  is small, however, the reversals are less distinct. It is less certain whether a given trade is a buy or sell. The allocation of the price change between the transient (bid-ask) component and the permanent change in the security value is less clear. This naturally leads to greater uncertainty (less concentration) and the negative correlation (downward slope) implied by the posterior in Panel A.

This illustration has implications for studies of U.S. equities. Although prior to 2000 the minimum price increment on most U.S. equities was \$0.125, it has since been \$0.01, and currently this value might well approximate the posted half-spread in a large, actively traded issue. For a share hypothetically priced at \$50, the implied  $c$  equals 0.0002. No approach using daily trade data is likely to achieve a precise estimate of such a magnitude. The posted half-spread for a thinly traded issue might be 25 cents on a \$5 stock, implying  $c$  equals 0.05. This is likely to be estimated much more precisely.

### *E. Small Sample Properties and Other Practical Considerations*

As emphasized above, the effective cost parameter  $c$  in this model is a regression coefficient. In the standard Bayesian normal regression framework, the coefficient posterior distribution is a combination of the prior and the distribution of the conventional OLS coefficient estimate. Drawing on the usual Gauss–Markov results, the latter estimate is unbiased. In a large data-dominated sample, therefore, any bias in the posterior estimate of  $c$  should be small. In small samples, however, the posterior will more closely resemble the prior. This is important because the prior is often strongly biased, due to the nonnegativity restriction. The mean of the prior used in the illustration of the last section,  $N^+(0, 1)$ , is  $\sqrt{2/\pi} \approx 0.8$ , a value much higher than a plausible  $c$  for any U.S. stock. The mean of the prior used in the actual implementations,  $N^+(0, 0.05^2)$ , is approximately 0.04. While this might be close to the mean  $c$  in a cross-section of U.S. stocks, the range is likely to be large, with values for some individual firms far removed from 0.04.

The usual way of characterizing variation in a liquidity parameter is to construct a series of estimates based on short samples. Common practice uses monthly estimates for an individual firm based on daily data, that is, roughly 20 observations. In the present situation, however, the posterior for a sample this size will closely resemble the prior. As a result, the monthly estimates for individual stocks will generally be highly biased. The relatively poor performance of monthly Gibbs estimates is noted by Goyenko et al. (2005). This bias will extend to subsequently derived estimates of other parameters (like systematic variability). As the bias will tend to be in the same direction for similar stocks, portfolio formation is unlikely to mitigate the problem.

These considerations do not rule out the use of the present approach to characterize variation in  $c$ , however. The problems arise from the practice of constructing a dynamic series by estimating over progressively smaller time windows. Section VI discusses alternative approaches to capturing variation.

It was noted above that one of the advantages of the Gibbs approach is that it restricts the effective cost estimate to be positive, thereby avoiding the infeasibility problem that arises in applying the moment estimate when the price change autocovariance is (in sample) positive. This is an important difference, but it is nevertheless clear from the illustration in the last section that the feature of the data driving the Gibbs estimate is the prominence and prevalence of price reversals. Since these reversals will also tend to generate the negative

Table I  
Variable Definitions

Effective cost measures	$c_{it}^{TAQ}$	Estimate for firm $i$ in year $t$ based on high-frequency (TAQ) data. For a given trade, the effective cost is the difference between the log transaction price and the prevailing log quote midpoint. $c_{it}^{TAQ}$ is the average over all trades in the year, weighted by dollar value of the trade.
	$c_{it}^{Gibbs}$	Gibbs estimate for firm $i$ in year $t$ using the market-factor model applied to daily CRSP prices and dividends.
	$c_{it}^{Moment}$	Modified Roll estimate for firm $i$ in year $t$ , equal to $\sqrt{-\text{Cov}(\Delta p_t, \Delta p_{t-1})}$ where $\Delta p_t$ is the log price change and the autocovariance, $\text{Cov}(\Delta p_t, \Delta p_{t-1})$ is estimated over all trading days in the year. $c_{it}^{Moment}$ is set to zero if the autocovariance is positive.
Other liquidity measures	$\lambda_{it}$	Price impact coefficient for firm $i$ in year $t$ based on TAQ data and estimated from the regression $\Delta p_\tau = \lambda(\text{Signed}\sqrt{\text{Dollar Volume}})_\tau + \varepsilon_\tau$ estimated annually using log price changes and aggregated signed dollar volumes where $\tau$ indexes five-minute intervals.
	$I_{it}$	The Amihud (2002) illiquidity measure for firm $i$ in year $t$ , $ daily\ return / daily\ dollar\ volume $ , averaged over all days with nonzero volume.
	$PropZero_{it}$	Proportion of trading days in the year that had a zero price change from the previous day, for firm $i$ in year $t$ .
Variables used in asset-pricing specifications	$r_{mt}$	Return on valued-weighted portfolio of NYSE, Amex, and NASDAQ stocks (Fama–French).
	$r_{ft}$	Return on 1-month T-bills (Ibbotson Associates) in month $t$ .
	$R_{it}$	Excess return, $r_{it} - r_{mt}$ , on portfolio $i$ in month $t$ .
	$smb_t$	Fama–French size factor in month $t$ .
	$hml_t$	Fama–French book-to-market factor in month $t$ .
	$\beta_{m,it}^{Gibbs}$	Market beta for portfolio $i$ in month $t$ . (Portfolio average of Gibbs estimates for prior year.)
	$LRMC_{it}$	Log relative market capitalization for portfolio $i$ in month $t$ . Let $m_{jt}$ denote the logarithm of the equity market capitalization of firm $j$ at the end of preceding year, $LRMC_{jt} = m_{jt} - \text{median}(m_{kt})$ where the median is computed over all firms in the sample at the end of the previous year. $LRMC_{it}$ is the average over all firms in portfolio $i$ .

autocovariance driving the Roll estimate, there is a similarity in the way the two estimators exploit the data.

Knowing the feature of the data that is driving the estimates helps us to think about other economic mechanisms that might be affecting the estimates. Among other things, price reversals can be generated (at various horizons) by market makers’ dynamic inventory control (Amihud and Mendelson (1980), Hasbrouck and Sofianos (1993), Madhavan and Smidt (1993)); by changing risk aversion (Campbell, Grossman, and Wang (1993)); by changing exposure to the risk of nonmarketable wealth (e.g., Lo, Mamaysky, and Wang (2004)); etc. Indeed, virtually any mechanism that features stationary variation in preferences, cash flows, information, or irrational trading can in principle induce price reversals.

Realistic calibration, however, often suggests that the magnitudes of these effects, at least at the daily frequency, are likely to be small.

## II. Data and Implementation

### A. Sample Construction

Most of the Gibbs estimates in the paper are computed in annual samples of daily data. These data are taken from the 1926 to 2006 CRSP daily data set, restricted to ordinary common shares (CRSP share code 10 or 11) that had a valid price for the last trading day of the year, and had no changes of listing venue or large splits within the last 3 months of the year. For purposes of assessing the performance of the Gibbs estimates, the analysis uses TAQ data produced by the NYSE. The asset pricing tests also use the Fama–French return factors (downloaded from Ken French’s web site).

Although the full CRSP sample is used in the asset pricing tests, the performance of the Gibbs estimates is assessed using a smaller comparison sample. This sample consists of 300 randomly chosen firms per year, 1993 to 2005. Liquidity measures for these firms are estimated from the TAQ data set. These 3,900 CRSP firm-years are matched to TAQ subject to the criteria of: agreement of ticker symbol; uniqueness of ticker symbol; the correlation (over the year) between the TAQ and CRSP closing prices is above 0.9; and, on fewer than 2% of the days does TAQ report trades when CRSP does not (or vice versa). Subject to these criteria, 3,777 firms are matched between TAQ and CRSP. Table II reports summary statistics for the comparison sample.

### B. TAQ Liquidity Measures

In the comparison sample, the effective cost for firm  $i$  on day  $t$  is computed as a trade-weighted average for all trades relative to the prevailing quote midpoint.

**Table II**  
**Summary Statistics for the Comparison Sample, 1993 to 2005**

The comparison sample consists of approximately 150 NASDAQ firms and 150 NYSE/Amex firms selected in a capitalization-stratified random draw in each of the years from 1993 to 2005. Values in the table are based on annual estimates for the 3,777 firms that could be matched between CRSP and TAQ. Variable definitions are given in Table I.

Variable	Mean	Median	Std. Dev.	Skewness	Kurtosis
$c_{it}^{TAQ}$	0.0106	0.0054	0.0146	4.61	54.7
$c_{it}^{Gibbs}$	0.0112	0.0061	0.0141	4.97	62.8
$c_{it}^{Moment}$	0.0106	0.0056	0.0152	4.35	52.1
$PropZero_{it}$	0.1363	0.1071	0.1171	1.02	0.9
$\lambda_{it} \times 10^6$	28.1500	7.4098	70.6173	7.84	101.2
$I_{it} \times 10^6$	3.6592	0.0709	20.0366	16.56	395.8
Market capitalization (\$ Million)	2,587.7190	196.9200	14,407.3199	18.55	502.9
Price (end of year, \$/share)	20.8442	14.5000	29.4357	11.38	229.8

Similar results obtain using unweighted averages.<sup>3</sup> In principle, the effective cost measures the cost of an order executed as a single trade. When the order is executed in multiple trades, the price impact of a trade also contributes to the execution cost. For each firm in the comparison sample, a representative price impact coefficient is estimated as the  $\lambda_i$  coefficient in the regression

$$\Delta p_{it} = \lambda_i (\text{Signed } \sqrt{\text{Dollar Volume}})_{it} + \varepsilon_{it}. \quad (4)$$

The specification is estimated using price changes and signed volume aggregated over 5-minute intervals. A separate estimate is computed for each month. Reported summary statistics are based on the average of the monthly values. Variants of specification (4) were also employed, with qualitatively similar results.

### *C. CRSP Liquidity Measures*

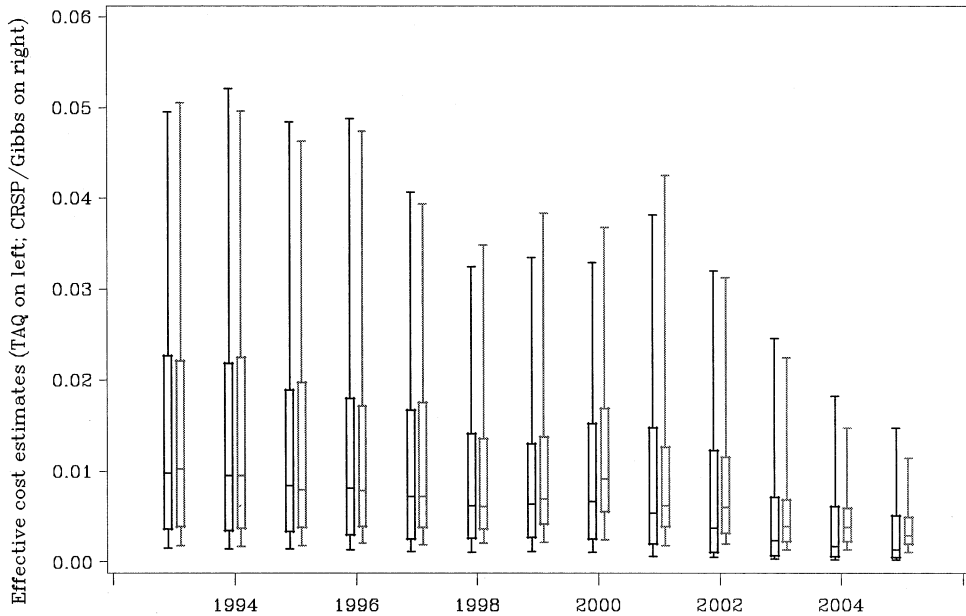
The study considers various alternative daily liquidity proxies. The simplest is the moment estimate of the effective cost based on the traditional Roll model, that is,  $\sqrt{-\text{Cov}(\Delta p_{i,t}, \Delta p_{i,t-1})}$ . When the autocovariance is positive, the moment estimate is set to zero. (This occurs for roughly one-third of the firm-years in the comparison sample.) The statistics reported in the paper use only those days on which trading occurred, but similar results are obtained when all prices (including nontrade days) are used.

In addition, the analysis includes the proportion of days with no price changes relative to the previous close (Lesmond, Ogden, and Trzcinka (1999)) and the Amihud (2002) illiquidity measure ( $I = |\text{return}|/|\text{Dollar volume}|$ ). The study does not include the Pástor and Stambaugh (2003) gamma measure because the authors caution against its use as a liquidity measure for individual stocks, noting the large sampling error in the individual estimates (p. 679).

## **III. Results in the Comparison Sample**

Table II presents summary statistics for the TAQ and CRSP liquidity variables. Since the effective costs are logarithmic, the means correspond to effective costs of about 1%. The proportion of zero returns is restricted to the unit interval by construction. At its median value, the TAQ-based price impact coefficient  $\lambda$  implies that a \$10,000 buy order would move the log price by  $\sqrt{10,000} \times 7 \times 10^{-6} = 0.0007$ , that is, seven basis points. The median value for the illiquidity ratio suggests that \$10,000 of daily volume would move the price by  $10,000 \times 0.07 \times 10^{-6} = 0.0007$  as well. The summary statistics of both the CRSP moment and Gibbs estimates of effective costs are close to the

<sup>3</sup> The prevailing quote is assumed to be the most recent quote posted two seconds or more prior to the trade. This is within the "1 to 2 seconds" rule that Piowar and Wei (2006) find optimal for their 1999 sample, but it is likely that reporting conventions have changed over the sample period used here.



**Figure 2. TAQ and CRSP/Gibbs estimates of effective cost in the comparison sample.** The comparison sample consists of approximately 150 NASDAQ firms and 150 NYSE/Amex firms selected in a capitalization-stratified random draw in each of the years 1993 to 2005. For each firm in each year, the effective cost is estimated from TAQ data and from CRSP daily data using the Gibbs procedure. The figure depicts the cross-sectional distributions for these estimates year-by-year, with TAQ estimates on the left and Gibbs estimates on the right. The upper and lower ranges of the box-and-whisker figures demarcate the 5<sup>th</sup> and 95<sup>th</sup> percentiles; the upper and lower edges of the boxes correspond to the 25<sup>th</sup> and 75<sup>th</sup> percentiles; the line drawn across the box indicates the median.

TAQ values. All liquidity measures exhibit extreme values; the coefficients of skewness and kurtosis are large, particularly for the illiquidity measure.

The discussion now focuses more closely on effective costs. Figure 2 presents annual box-and-whisker plots of the TAQ and CRSP/Gibbs estimates. There are several notable features of the TAQ values. First, the distributions do not appear stationary. Although the 5<sup>th</sup> percentile (indicated by the lower limit of the whisker) is relatively stable, the 95<sup>th</sup> percentile (upper limit of the whisker) has declined from about 0.05 in 1993 to 0.02 in 2005. The median (marked by the horizontal line in the box) has declined from roughly 0.01 in 1993 to 0.004 in 2005. This decline may reflect changes in trading technology and regulation, but it may also arise from changes in the composition of the sample.

The second important feature is that cross-sectional variation is larger than the aggregate time-series variation. The smallest range between the 5<sup>th</sup> and 95<sup>th</sup> percentiles is about 0.01 (in 2005), and for most of the sample the range is at least 0.02. This dominates the decline in the median over the period, roughly 0.006. This suggests that tests of liquidity effects are likely to be more powerful if they are based on cross-sectional variation.

Table III  
Correlations between Liquidity Measures for the Comparison Sample

The comparison sample consists of approximately 150 NASDAQ firms and 150 NYSE/Amex firms selected in a capitalization-stratified random draw in each of the years from 1993 to 2005. Definitions of the liquidity measures are given in Table I. Partial correlations are adjusted for log (end-of-year price) and log (market capitalization).

	$c_{it}^{TAQ}$	$c_{it}^{Gibbs}$	$c_{it}^{Moment}$	$PropZero_{it}$	$\lambda_{it}$	$I_{it}$
Pearson correlation						
$c_{it}^{TAQ}$	1.000	0.965	0.878	0.611	0.513	0.612
$c_{it}^{Gibbs}$	0.965	1.000	0.917	0.579	0.450	0.589
$c_{it}^{Moment}$	0.878	0.917	1.000	0.451	0.378	0.504
$PropZero_{it}$	0.611	0.579	0.451	1.000	0.311	0.252
$\lambda_{it}$	0.513	0.450	0.378	0.311	1.000	0.668
$I_{it}$	0.612	0.589	0.504	0.252	0.668	1.000
Spearman correlation						
$c_{it}^{TAQ}$	1.000	0.872	0.636	0.770	0.735	0.937
$c_{it}^{Gibbs}$	0.872	1.000	0.791	0.620	0.577	0.778
$c_{it}^{Moment}$	0.636	0.791	1.000	0.417	0.363	0.592
$PropZero_{it}$	0.770	0.620	0.417	1.000	0.510	0.704
$\lambda_{it}$	0.735	0.577	0.363	0.510	1.000	0.824
$I_{it}$	0.937	0.778	0.592	0.704	0.824	1.000
Pearson partial correlation						
$c_{it}^{TAQ}$	1.000	0.943	0.805	0.366	0.268	0.567
$c_{it}^{Gibbs}$	0.943	1.000	0.866	0.359	0.189	0.517
$c_{it}^{Moment}$	0.805	0.866	1.000	0.193	0.107	0.397
$PropZero_{it}$	0.366	0.359	0.193	1.000	0.068	0.103
$\lambda_{it}$	0.268	0.189	0.107	0.068	1.000	0.610
$I_{it}$	0.567	0.517	0.397	0.103	0.610	1.000
Spearman partial correlation						
$c_{it}^{TAQ}$	1.000	0.678	0.382	0.564	0.024	0.631
$c_{it}^{Gibbs}$	0.678	1.000	0.682	0.285	-0.123	0.361
$c_{it}^{Moment}$	0.382	0.682	1.000	0.101	-0.182	0.288
$PropZero_{it}$	0.564	0.285	0.101	1.000	-0.021	0.341
$\lambda_{it}$	0.024	-0.123	-0.182	-0.021	1.000	0.375
$I_{it}$	0.631	0.361	0.288	0.341	0.375	1.000

The general features of the CRSP/Gibbs cost distributions closely match those derived from TAQ. To more rigorously assess the quality of the CRSP/Gibbs estimates and other liquidity proxies, Table III presents the correlation coefficients. The standard (Pearson) correlation between the TAQ and CRSP/Gibbs estimate of effective cost is 0.965.<sup>4</sup> The Spearman correlation, a more appropriate

<sup>4</sup> This and other reported correlations are computed as a single estimate, pooled over years and firms. The values are very similar, though, to the averages of annual cross-sectional correlations. Over the 13-year sample, the lowest estimated correlation between the CRSP/Gibbs estimate and the TAQ value is 0.903 (in 2005, possibly reflecting the narrowing of spreads postdecimalization).

measure if the proxy is being used to rank liquidity, is 0.872. Because liquidity proxies are often used in specifications with explanatory variables that are themselves likely to be correlated with liquidity, the table also presents partial correlations that control for the logarithm of end-of-year share price and the logarithm of market capitalization. Not only are the CRSP/Gibbs estimates strong proxies in the sense of correlation, but they are also good point estimates of the TAQ measures. A regression of the latter against the former would ideally have unit slope and zero intercept. In the comparison sample, the estimated regression is  $c_i^{TAQ} = 0.001 + 0.935c_i^{CRSP/Gibbs} + e_i$ . By any of the four types of correlation considered here, the conventional moment estimate of effective cost is dominated by the CRSP/Gibbs estimator.

The table also reports correlations for the alternative TAQ and CRSP liquidity measures. The two TAQ-based liquidity measures (effective cost and price impact coefficient) are moderately positively correlated (0.513, Pearson). This is qualitatively similar to the findings of Korajczyk and Sadka (2008). Among the daily proxies, the Amihud illiquidity measure is most strongly correlated with the TAQ-based price impact coefficient, with the CRSP/Gibbs effective cost estimate being second.

#### IV. Historical Estimates, 1926 to 2006

##### A. Effective Cost

The basic market-factor model is estimated annually for all ordinary common shares in the CRSP daily database. Figure 3 graphs effective costs, separately for NYSE/Amex (listed) and NASDAQ, averaged over market capitalization quartiles.

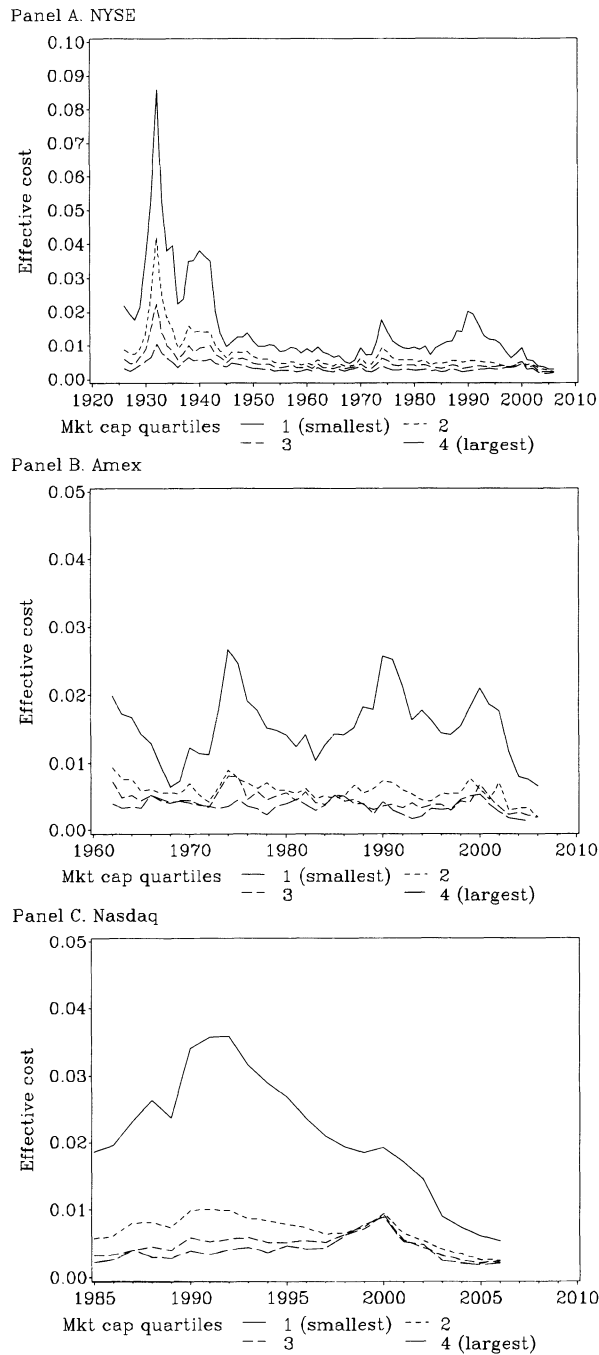
Effective costs for NYSE/Amex issues (upper graph) exhibit considerable variation over time. The highest values are found immediately after the 1929 crash and during the Depression. It is likely that this reflects historic lows for per-share prices coupled with a tick size that remained at one-eighth of a dollar, which together imply an elevated proportional cost. Subsequent peaks in effective cost generally also coincide with local minima of per-share prices. After the Depression, however, average effective costs don't rise above 1% for the three highest capitalization quartiles. The largest variation is confined to the bottom capitalization quartile.

The NASDAQ estimates (lower graph) begin in 1985. As for the listed sample, the largest variation arises in the lowest capitalization quartile. The temporal variation, however, may also reflect changes in sample composition. In the early 1990s, NASDAQ delisted firms that were especially young and volatile (Fama and French (2004), Fink et al. (2006)).

##### B. Trade Directions

Although the discussion has emphasized the estimates of model parameters, the Gibbs procedure also generates posteriors for the trade direction indicators (the  $q_t$ ). These offer insight into the model because they help assess the validity





**Figure 3. Average effective costs 1926 to 2006.** Average Gibbs effective cost estimates for all ordinary common shares in the CRSP daily database. NYSE, Amex, and NASDAQ firms are analyzed separately; subsamples are quartiles based on end-of-year market capitalization. Fama-French NYSE breakpoints are used to construct the samples.

of the assumptions and suggest ways in which the model might be extended. They also have implications for broader phenomena. The returns usually used in asset pricing specifications are based on last-trade prices, and therefore reflect bid/ask components that are driven by these indicators. Any commonality or seasonality in these indicators is likely therefore to contribute to commonality and seasonality in returns.

The analyses are based on the set of  $\hat{q}_{it}$ , which denotes the estimated posterior mean of the trade direction indicator for firm  $i$  on day  $t$ . There are roughly 22,000 firms and 21,520 days (all trading days, 1926 to 2006), but most firms are traded only for a portion of the sample. The average of  $\hat{q}_{it}$  over all firms and days is, at  $-0.008$ , quite close to zero. Not surprisingly, however, given the large number of observations, the hypothesis of a zero mean is easily rejected. In principle the  $q_t$  in the Roll model have unit variance. The standard deviation of the  $\hat{q}_{it}$ , however, is 0.379 (a variance of 0.143). These lower values arise because the  $\hat{q}_{it}$  are posterior estimates. The prior mean is zero, and the sample is rarely sufficiently informative to confidently assert that a particular trade is a buy or a sale. Furthermore, following the CRSP midpoint convention discussed in Section I.C, some of the trade direction indicators are set to zero.

In the development of the sampler, the  $q_t$  are assumed to be serially uncorrelated. Over the entire sample, the average first-order autocorrelation,  $\text{Corr}(\hat{q}_{it}, \hat{q}_{i,t-1})$ , is  $-0.32$ . The violation of the assumption is not as large, though, as it might first appear. The autocorrelation is computed as the autocovariance divided by the variance, and as noted above, the variance of the  $\hat{q}_{it}$  is much lower than that of the true underlying measures. The average first-order autocovariance  $\text{Cov}(\hat{q}_{it}, \hat{q}_{i,t-1})$  is, at  $-0.042$ , much closer to zero.

The  $\Delta q_t$  in the market-factor model (3) are also assumed to be uncorrelated with  $r_{mt}$ . Across all firms, the  $\text{Corr}(\Delta \hat{q}_{it}, r_{mt})$  are generally close to zero: The 1<sup>st</sup>, 50<sup>th</sup>, and 99<sup>th</sup> percentiles are (respectively)  $-0.091$ ,  $-0.004$ , and  $0.072$ . However, the  $\hat{q}_{it}$  estimates do not offer any insight into the appropriateness of the assumption for specifications (2) and (3) that  $\text{Corr}(\Delta q_t, u_t) = 0$ . This is because in each sweep of the sampler, these specifications are estimated via OLS, and the computed residuals (the  $u_t$ ) are orthogonal to the dependent variables (the  $\Delta q_t$ ) by construction.

The  $\hat{q}_{it}$  estimates may be used to assess cross-firm commonality in trade directions. Let  $\bar{q}_t$  denote the cross-firm average of the  $\hat{q}_{it}$  on day  $t$ , with  $t = 1, \dots, 21,520$  (all of the trading days, 1926 to 2006). The 10<sup>th</sup> and 90<sup>th</sup> percentiles of the  $\bar{q}_t$  are  $-0.047$  and  $0.029$ , respectively, suggesting that on any given day there are modest systematic cross-firm patterns in trade directions. More formally, I compute for each day the  $p$ -value for the null hypothesis that  $E[\bar{q}_t] = 0$ , assuming independence of the  $\hat{q}_{it}$  across firms. Roughly half of these  $p$ -values are below 0.03. The commonality may also be characterized by examining the correlations  $\text{Corr}(\hat{q}_{it}, \bar{q}_t)$  estimated for each firm over all days with observations. The average (across all firms)  $\text{Corr}(\hat{q}_{it}, \bar{q}_t)$  is 0.093. The  $t$ -statistic for the null hypothesis of a zero mean, assuming independence, is 248.

The return specifications discussed later in the paper will be seen to exhibit striking liquidity-related seasonalities. In this connection, it is useful to

consider seasonalities in the  $\hat{q}_{it}$ . The return specifications are estimated using monthly data, which naturally suggests consideration of the trade direction of the last price of the month. Figure 4 depicts (by exchange) the averages of these estimates.

Common to all exchanges is a pattern of end-of-quarter and end-of-year elevations. This implies that end-of-quarter trade prices are more likely to be at the ask price, a finding consistent with institutional window dressing (Lakonishok, Thaler, and Vishny (1991), Musto (1997), Sias and Starks (1997), Musto (1999), O'Neal (2001), He, Ng, and Wang (2004), Meier and Schaumburg (2004), Elton et al. (2006), and Sias (2006)).

## V. Stock Returns and Effective Cost

This section presents empirical analyses aimed at determining whether the level of effective cost is a priced characteristic in long-term U.S. equity data.

### A. Specifications and Estimation Methodology

The empirical analysis follows the GMM approach summarized in Cochrane (2005) (pp. 241–243), modified to allow for characteristics and applied to portfolios constructed according to various rankings. The specification for expected returns is

$$ER_t = \beta\lambda + Z_t\delta, \quad (5)$$

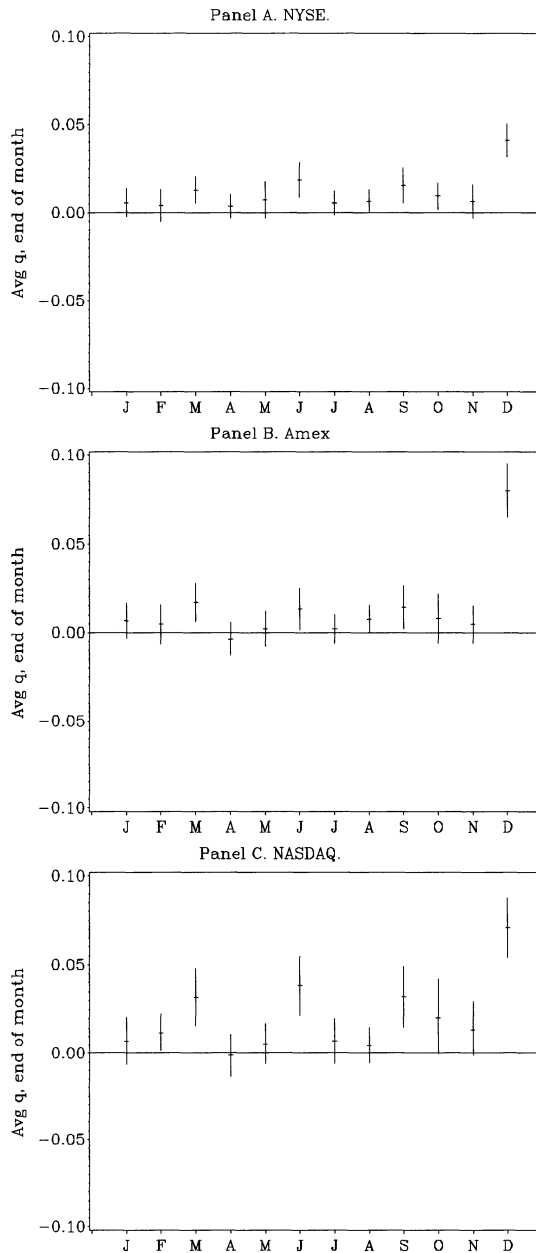
where  $R_t$  is a vector of excess returns relative to the risk-free rate for  $N$  assets;  $\lambda$  is a  $K$ -vector of factor risk premia;  $\beta$  is a matrix of factor loadings;  $Z_t$  is an  $N \times M$  matrix of characteristics; and  $\delta$  is an  $M$ -vector of coefficients for the characteristics. The factor loadings are the projection coefficients in the  $K$ -factor return generating process

$$R_t = a + \beta f_t + u_t, \quad (6)$$

where  $a$  is a constant vector;  $f_t$  is a vector of factor realizations; and  $u_t$  is a vector of idiosyncratic zero-mean disturbances. The equilibrium conditions that follow from the usual economic arguments imply  $\delta = 0$  and  $a = \beta(\lambda - Ef_t)$ . If all factors are excess returns on traded portfolios (a condition met for all factors used here) the second condition reduces to  $a = 0$ .

The results reported below are representative of a large set of potential specifications. Two sets of factors are considered. The first set consists solely of the Fama–French excess market return factor,  $r_{mt} - r_{ft}$ . The second set adds the  $smb_t$  and  $hml_t$  factors.

The characteristics are  $c^{Gibbs}$  (estimated in the prior year), relative size, and various January seasonal terms. The relative size measure is constructed as follows. Letting  $m_{jt}$  denote the logarithm of the equity market capitalization of firm  $j$  at the end of the preceding year, the log market capitalization relative to the median is  $m_{jt} - \text{median}(m_{kt})$  where the median is computed over



**Figure 4. Trade directions associated with end-of-month prices.** In the Roll model the trade direction variable  $q_t = +1$  if the trade is at the ask price,  $q_t = -1$  if at the bid, and  $q_t = 0$  if the reported price is quote midpoint. The Gibbs estimation procedure produces estimates of  $q_t$  for each reported price. The figure depicts the means of the  $q_t$  across firms and years for the last reported price (trade or quote midpoint) of the month. Means are indicated by a horizontal line. Standard errors of the means are computed by first grouping over firms (for a given month and year). The vertical lines demarcate the mean  $\pm$  twice the standard error. The sample covers 1926 to 2006 for NYSE firms, 1962 to 2006 for Amex, and 1985 to 2006 for NASDAQ.

all of the firms in the sample. The log relative market capitalization,  $LRMC_{it}$ , is then computed as the average over all firms in portfolio  $i$ . The normalization by median firm size captures the cross-sectional variation while removing the nonstationary long-run components. The seasonal variables are a January dummy ( $JanDum_t$ ) and January interactions with  $c^{Gibbs}$  and  $LRMC$ . As the characteristics are not demeaned,  $Z_t$  also includes a constant term. With these definitions, specification (5) becomes

$$\begin{aligned}
 R_{it} - r_{ft} = & \delta_0 + \lambda_m \beta_i^m + \lambda_{smb} \beta_i^{smb} + \lambda_{hml} \beta_i^{hml} + \delta_{Jan} JanDum_t \\
 & + \left[ \delta_c c_{it}^{Gibbs} \right. \\
 & \quad \left. + \delta_{c \times Jan} (c_{it}^{Gibbs} \times JanDum_t) + \delta_{c \times \sim Jan} (c_{it}^{Gibbs} \times (1 - JanDum_t)) \right] \\
 & + \delta_{LRMC \times Jan} (LRMC_{it} \times JanDum_t) \\
 & + \delta_{LRMC \times \sim Jan} (LRMC_{it} \times (1 - JanDum_t)) + u_{it}.
 \end{aligned} \tag{7}$$

Within the bracket, the top and bottom expressions are mutually exclusive (to avoid linear dependence).

The model is estimated using monthly return data and a GMM procedure that estimates (6) and (7) jointly (Cochrane (2005)). The parameter estimates reported below are equivalent to those obtained from a two-pass procedure in which estimates of  $\beta$  are obtained via OLS time-series regression of (6) over the full sample and then used on the right-hand side in an OLS estimation of (7). The GMM standard errors of the  $\lambda$  and  $\delta$  estimates are corrected for estimation error in the  $\beta$  values (as well as heteroskedasticity).<sup>5</sup>

### B. Portfolio Formation

Portfolios are formed annually based on information available at the start of the year: market capitalization at the close of the prior year, and the Gibbs estimates of effective cost and beta formed over the prior year. In the return specifications, portfolio values are equally weighted averages.<sup>6</sup>

<sup>5</sup> More precisely, the moment conditions used in estimation are

$$E \begin{bmatrix} R_t - (a + \beta f_t) \\ f_t' \otimes (R_t - (a + \beta f_t)) \\ \beta'(R_t - \beta\lambda - Z_t\delta) \\ Z_t'(R_t - \beta\lambda - Z_t\delta) \end{bmatrix} = 0.$$

These suffice to identify estimates of  $a$ ,  $\beta$ ,  $\lambda$ , and  $\delta$  that equal those from the two-pass OLS procedure. The first two (vector) conditions are the  $N(K+1)$  normal equations that identify the estimates of  $a$  and  $\beta$ ; the second two conditions are the  $K+M$  normal equations that identify the estimates of  $\lambda$  and  $\delta$ . Cochrane shows that under the assumption of normality, the GMM standard errors are asymptotically equivalent to those constructed with the Shanken (1992) correction.

<sup>6</sup> Due to the inverse relationship between market capitalization and liquidity, averages weighted by market capitalization (value-weighted averages) tend to suppress variation in effective cost. Accordingly, in alternative specifications that use value-weighted averages for returns and effective costs, the effective cost estimates are similar to those in the equally weighted specifications, but statistically weaker.

Results are reported for four samples. The first sample is a benchmark analysis. There are 10 NYSE portfolios ranked by market capitalization, and the estimation period is 1927 to 1998. It is straightforward to construct and analyze, and estimates are reported in Cochrane (2005). This helps in gauging the effects of liquidity measures in a standard specification.

The other three samples are derived from NYSE, Amex, and NASDAQ issues, and each are analyzed over the longest available span. The portfolio construction maximizes the variation in beta and effective cost. For the NYSE, for example, all stocks in the sample at the end of a year are ranked in five groups by the beta estimate. Then, within each beta group, stocks are ranked by effective cost. In the portfolio formation process, both beta and effective cost estimates are the Gibbs estimates of the basic market-factor model. The NYSE sample covers 1927 to 2006. The Amex (1963 to 2006) and NASDAQ samples (1985 to 2006) are constructed in a similar fashion.

Table IV presents summary statistics for the factors discussed above and related series for the four samples. All three Fama–French factors have positive

**Table IV**  
**Summary Statistics for Return Factors**

Summary statistics for return factors (and the risk-free rate) over each of the four sample periods considered in the return specifications. Variable definitions are given in Table I.

	Mean	<i>SD</i>	First-Order Autocorrelation	Correlation with:			
				$r_f$	$r_m - r_f$	<i>smb</i>	<i>hml</i>
Panel A: Benchmark Sample (1927 to 1998)							
$r_f$	0.0031	0.0026	0.973	1.000	−0.069	−0.057	0.008
$r_m - r_f$	0.0069	0.0556	0.111	−0.069	1.000	0.344	0.291
<i>smb</i>	0.0020	0.0321	0.119	−0.057	0.344	1.000	0.206
<i>hml</i>	0.0039	0.0354	0.194	0.008	0.291	0.206	1.000
Panel B: NYSE Sample (1927 to 2006)							
$r_f$	0.0031	0.0025	0.973	1.000	−0.069	−0.059	0.012
$r_m - r_f$	0.0065	0.0545	0.107	−0.069	1.000	0.329	0.215
<i>smb</i>	0.0025	0.0336	0.073	−0.059	0.329	1.000	0.091
<i>hml</i>	0.0042	0.0360	0.176	0.012	0.215	0.091	1.000
Panel C: Amex Sample (1963 to 2006)							
$r_f$	0.0047	0.0022	0.952	1.000	−0.096	−0.073	0.024
$r_m - r_f$	0.0049	0.0437	0.053	−0.096	1.000	0.301	−0.409
<i>smb</i>	0.0025	0.0321	0.068	−0.073	0.301	1.000	−0.280
<i>hml</i>	0.0047	0.0289	0.133	0.024	−0.409	−0.280	1.000
Panel D: NASDAQ Sample (1985 to 2006)							
$r_f$	0.0039	0.0017	0.949	1.000	−0.001	−0.158	−0.075
$r_m - r_f$	0.0070	0.0435	0.043	−0.001	1.000	0.197	−0.497
<i>smb</i>	0.0006	0.0340	−0.033	−0.158	0.197	1.000	−0.433
<i>hml</i>	0.0036	0.0314	0.097	−0.075	−0.497	−0.433	1.000

average returns in all sample periods. The risk-free rate is the most persistent series. Table V reports summary statistics for a subset of the portfolios in each sample. Note that the effective cost in the highest quintile is sharply higher, relative to the lower quintiles. This is consistent with the positive skewness of effective costs noted in connection with Table II.

### C. Results

Table VI reports estimates of the expected return specifications. Panel A contains estimates for the benchmark sample. Specification (1) is a one-factor CAPM specification. The estimated market price of risk is 0.0080 (0.80% per month) with a heteroskedasticity-consistent  $t$ -statistic of 4.11. Cochrane reports an estimate of 0.71% per month, with a  $t$ -statistic of 3.74. The discrepancies are small, and presumably reflect minor differences in portfolio construction and a slightly different time period. (Cochrane's sample begins in 1926, whereas here the first year is used to estimate the model parameters.) Specification (6) also includes *smb* (size) and *hml* (book-to-market) factors. The prices of market and book-to-market factor risk are positive and significant, but that of the size factor is close to zero.

Specifications (3) and (8) add as characteristics  $c_{it}^{Gibbs}$  and a constant. In both cases, the  $c_{it}^{Gibbs}$  coefficient is positive (but only marginally significant). This is consistent with the usual view that securities are priced so that their gross returns include compensation for trading costs. The point estimates, however, indicate that this notion should not be embraced unreservedly. The  $c_{it}^{Gibbs}$  coefficients in both specifications are slightly above one. This magnitude can only be justified by assuming an unrealistically high turnover. To see this, note that following Amihud and Mendelson (1986), a representative trader making a round-trip over 2 months would pay the effective cost twice (once on the purchase and once on the sale). This suggests that the expected 2-month gross return should impound  $2c_{it}^{Gibbs}$ , or equivalently, that the coefficient on  $c_{it}^{Gibbs}$  determining the expected 1-month gross return should be one. A holding period of 2 months is equivalent to a turnover of six times per year. While there is certainly considerable variation over time, firms, and investors, average annual NYSE turnover rates are currently around one, and are well under one for most of the 20<sup>th</sup> century. (The historical data section of the NYSE web site reports 2005 turnover as 103%; 1960 turnover was 12%.) Thus, the point estimates of the coefficient are difficult to reconcile with the most straightforward trading stories.

Specifications (4) and (9) introduce a January dummy and also split the effective cost variable into January and non-January components. The coefficients on January effective cost are positive and significant, while those on the non-January components become marginally significant. The January seasonal is generally thought to be associated with firm size. In specifications that employ the relative size measure in lieu of effective cost, the coefficient on the January relative size measure is strongly negative, implying that smaller firms have elevated January returns. These results confirm the findings of numerous other researchers, and are not shown (for the sake of brevity).

Table V  
Summary Statistics for Portfolios

For the NYSE, Amex, and NASDAQ, portfolios are formed by the criteria indicated below. For a representative set of the portfolios, the table reports averages (over time) of the equally weighted portfolio averages. Variable definitions are given in Table I.

Panel A: Benchmark Sample, 1927 to 1998. Ten NYSE Portfolios Sorted by Market Capitalization

Capitalization Rank	$R_{it}$	$\beta_{it}^{Gibbs}$	$c_{it}^{Gibbs}$	$LRMC_{it}$	Number of Firms
1	0.017	0.872	0.023	-2.529	104
5	0.009	0.995	0.007	-0.204	107
10	0.007	1.014	0.003	2.944	109

Panel B: NYSE Sample, 1927 to 2006. Twenty-Five Portfolios Formed by Ranking First on  $\beta_m^{Gibbs}$ , and Then on  $c^{Gibbs}$

$\beta_m^{Gibbs}$ Rank	$c^{Gibbs}$ Rank	$R_{it}$	$\beta_{it}^{Gibbs}$	$c_{it}^{Gibbs}$	$LRMC_{it}$	Number of Firms
1	1	0.007	0.348	0.002	0.899	44
1	3	0.008	0.330	0.006	-0.077	45
1	5	0.016	0.290	0.022	-1.822	43
3	1	0.008	0.908	0.002	1.656	45
3	3	0.010	0.909	0.005	0.253	45
3	5	0.014	0.908	0.019	-1.496	43
5	1	0.007	1.659	0.003	1.540	44
5	3	0.009	1.747	0.006	0.233	45
5	5	0.014	1.741	0.015	-1.063	43

Panel C: Amex Sample, 1963 to 2006. Twenty-Five Portfolios Formed by Ranking First on  $\beta_m^{Gibbs}$ , and Then on  $c^{Gibbs}$

$\beta_m^{Gibbs}$ Rank	$c^{Gibbs}$ Rank	$R_{it}$	$\beta_{it}^{Gibbs}$	$c_{it}^{Gibbs}$	$LRMC_{it}$	Number of Firms
1	1	0.008	0.134	0.003	0.823	27
1	3	0.009	0.111	0.010	-0.270	27
1	5	0.026	0.045	0.038	-1.327	26
3	1	0.009	0.640	0.003	1.208	27
3	3	0.009	0.640	0.011	0.009	28
3	5	0.018	0.644	0.035	-1.083	27
5	1	0.002	1.509	0.004	1.790	27
5	3	0.003	1.493	0.010	0.640	27
5	5	0.020	1.403	0.033	-0.710	26

Panel D: NASDAQ Sample, 1985 to 2006. Twenty-Five Portfolios Formed by Ranking First on  $\beta_m^{Gibbs}$ , and Then on  $c^{Gibbs}$

$\beta_m^{Gibbs}$ Rank	$c^{Gibbs}$ Rank	$R_{it}$	$\beta_{it}^{Gibbs}$	$c_{it}^{Gibbs}$	$LRMC_{it}$	Number of Firms
1	1	0.009	0.092	0.008	0.462	118
1	3	0.009	0.042	0.020	-0.688	121
1	5	0.020	-0.018	0.050	-1.889	115
3	1	0.008	0.635	0.004	1.774	120
3	3	0.009	0.630	0.014	0.092	121
3	5	0.021	0.625	0.044	-1.686	115
5	1	0.006	1.619	0.004	2.294	121
5	3	0.003	1.633	0.009	1.151	122
5	5	0.008	1.499	0.029	-0.504	117



Table VI  
Expected Return Estimates

For the indicated set of portfolios, Fama–French factor loadings ( $\beta$ s) are estimated via

$$R_{it} \equiv r_{it} - r_{ft} = \alpha_i + \beta_i^m r_{mt} + \beta_i^{smb} smb_t + \beta_i^{hml} hml_t + e_{it},$$

where  $r_{it}$  is the equally weighted monthly return on portfolio  $i$  and other variables are defined in Table I. The estimated  $\beta$ s are then used in panel return specifications

$$R_{it} = \delta_0 + \lambda_m \beta_i^m + \lambda_{smb} \beta_i^{smb} + \lambda_{hml} \beta_i^{hml} + \delta_{Jan} JanDum_t + \left[ \delta_c c_{it}^{Gibbs} + \delta_{c \times Jan} (c_{it}^{Gibbs} \times JanDum_t) + \delta_{c \times \sim Jan} (c_{it}^{Gibbs} \times (1 - JanDum_t)) \right] + \delta_{LMC \times Jan} (LMC_{it} \times JanDum_t) + \delta_{LMC \times \sim Jan} (LMC_{it} \times (1 - JanDum_t)) + u_{it},$$

where  $JanDum_t$  is a dummy variable equal to one for January returns. The two terms in the brackets are mutually exclusive, to avoid multicollinearity.  $T$ -statistics computed using GMM standard errors that correct for estimation error in the  $\beta$ s and heteroskedasticity are reported in parentheses.

Parameter	Specification									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\delta_0$		-0.0094 (-2.21)	0.0298 (1.82)	0.0289 (1.74)	-0.0144 (-0.26)		0.0166 (1.59)	-0.0328 (-1.18)	-0.0338 (-1.20)	-0.2146 (-1.40)
$\lambda_m$	0.0080 (4.18)	0.0153 (3.87)	-0.0240 (-1.55)	-0.0238 (-1.52)	0.0093 (0.21)	0.0059 (3.05)	-0.0094 (-0.91)	0.0351 (1.34)	0.0353 (1.34)	0.1807 (1.42)
$\lambda_{smb}$						-0.0003 (-0.14)	0.0015 (0.87)	-0.0030 (-1.05)	-0.0030 (-1.03)	0.0410 (1.30)
$\lambda_{hml}$						0.0102 (2.72)	0.0065 (1.75)	-0.0103 (-1.02)	-0.0103 (-1.01)	-0.0285 (-1.01)
$\delta_c$			1.3059 (1.92)					1.3733 (1.92)		
$\delta_{Jan}$				0.0097 (0.87)	0.0186 (1.29)				0.0097 (0.86)	0.0186 (1.27)
$\delta_{c \times Jan}$				4.6096 (2.98)	3.7212 (2.06)				4.6779 (2.98)	3.9220 (2.09)

(continued)

Table VI—Continued

Panel A: Benchmark Sample, 1927 to 1998. Ten NYSE Portfolios Sorted by Market Capitalization									
Specification									
Parameter	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9) (10)
$\delta_{c \times \sim Jan}$				0.9943 (1.40)	1.0913 (1.50)			1.0619 (1.42)	1.2858 (1.51)
$\delta_{LRMC \times Jan}$					-0.0062 (-0.98)				0.0044 (0.40)
$\delta_{LRMC \times \sim Jan}$					0.0046 (0.98)				0.0151 (1.45)

Panel B: NYSE Sample, 1927 to 2006. Twenty-Five Portfolios Formed by Ranking First on $\beta_{Gibbs}^m$ , and Then on $c^{Gibbs}$									
Specification									
Parameter	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9) (10)
$\delta_0$		0.0047 (2.72)	0.0071 (2.70)	0.0061 (2.28)	0.0003 (0.07)		0.0070 (4.97)	0.0036 (1.36)	0.0026 (0.97)
$\lambda_m$	0.0078 (4.35)	0.0043 (1.65)	-0.0035 (-0.91)	-0.0035 (-0.89)	-0.0009 (-0.41)	0.0046 (1.63)	-0.0011 (-0.47)	0.0036 (1.00)	0.0025 (0.78)
$\lambda_{smb}$						0.0008 (0.07)	0.0013 (0.63)	-0.0053 (-1.14)	0.0001 (0.00)
$\lambda_{hml}$						0.0094 (0.56)	0.0074 (2.32)	-0.0070 (-0.87)	-0.0092 (-0.82)
$\delta_c$			0.9297 (2.07)					1.1446 (1.94)	
$\delta_{Jan}$				0.0121 (1.33)	0.0186 (1.49)			0.0121 (1.32)	0.0186 (1.48)
$\delta_{c \times Jan}$				3.6642 (3.02)	3.2202 (2.03)			3.8773 (3.04)	3.2411 (2.03)
$\delta_{c \times \sim Jan}$				0.6740 (1.41)	0.9602 (1.45)			0.8880 (1.43)	0.9815 (1.46)
$\delta_{LRMC \times Jan}$					-0.0062 (-0.93)				-0.0078 (-1.19)
$\delta_{LRMC \times \sim Jan}$					0.0045 (1.42)				0.0029 (1.01)

(continued)

Table VI—Continued

Panel C: Amex Sample, 1963 to 2006. Twenty-Five Portfolios Formed by Ranking on $\beta_m^{Gibbs}$ , and Then on $c^{Gibbs}$									
Parameter	Specification								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\delta_0$		0.0137 (4.12)	0.0136 (4.66)	0.0118 (4.15)	0.0138 (4.20)		0.0078 (1.88)	0.0166 (3.01)	0.0145 (2.67)
$\lambda_m$	0.0084 (3.26)	-0.0031 (-0.75)	-0.0131 (-3.56)	-0.0129 (-3.63)	-0.0208 (-4.00)	-0.0059 (-1.74)	-0.0138 (-2.60)	0.0056 (0.68)	0.0052 (0.68)
$\lambda_{smb}$						0.0079 (2.49)	0.0095 (2.95)	-0.0251 (-3.15)	-0.0242 (-3.46)
$\lambda_{hml}$						0.0129 (4.91)	0.0095 (2.94)	-0.0040 (-0.65)	-0.0036 (-0.62)
$\delta_c$			0.7939 (4.71)					1.4107 (3.56)	
$\delta_{Jan}$				0.0216 (2.04)	0.0073 (0.40)				0.0216 (2.06)
$\delta_{c \times Jan}$				4.7652 (6.36)	6.0935 (4.39)				5.3472 (6.31)
$\delta_{c \times \sim Jan}$				0.4139 (3.22)	0.8474 (3.14)				1.0060 (3.26)
$\delta_{LRMC \times Jan}$					0.0238 (1.78)				0.0191 (1.48)
$\delta_{LRMC \times \sim Jan}$					0.0073 (2.59)				0.0026 (1.20)

(continued)

Table VI—Continued

Panel D: NASDAQ Sample, 1985 to 2006. Twenty-Five Portfolios Formed by Ranking on  $\beta_m^{Gibbs}$ , and Then on  $c^{Gibbs}$

Parameter	Specification									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\delta_0$		0.0164 (6.26)	0.0091 (2.77)	0.0067 (1.88)	0.0063 (1.62)		0.0044 (0.51)	-0.0014 (-0.23)	-0.0036 (-0.63)	-0.0092 (-1.08)
$\lambda_m$	0.0069 (2.05)	-0.0058 (-1.44)	-0.0038 (-0.94)	-0.0039 (-0.96)	-0.0028 (-0.66)	-0.0025 (-0.59)	-0.0062 (-0.73)	0.0048 (0.55)	0.0042 (0.52)	0.0180 (1.08)
$\lambda_{smb}$						0.0121 (2.61)	0.0116 (2.70)	0.0009 (0.10)	0.0014 (0.17)	-0.0064 (-0.53)
$\lambda_{hml}$						0.0096 (3.41)	0.0073 (1.33)	0.0075 (2.04)	0.0074 (2.02)	0.0095 (1.67)
$\delta_c$			0.2727 (2.56)					0.2676 (1.64)		
$\delta_{Jan}$				0.0322 (1.76)	0.0391 (1.70)				0.0322 (1.76)	0.0391 (1.70)
$\delta_{c \times Jan}$				2.0292 (5.09)	1.6357 (2.10)				2.0187 (4.82)	1.6187 (2.07)
$\delta_{c \times \sim Jan}$				0.1014 (1.15)	0.0596 (0.33)				0.0911 (0.63)	0.0442 (0.24)
$\delta_{LRMC \times Jan}$					-0.0070 (-0.92)					-0.0088 (-1.12)
$\delta_{LRMC \times \sim Jan}$					-0.0008 (-0.38)					-0.0027 (-0.80)

Specifications (5) and (10) include a January dummy and January/non-January components of both effective cost and relative size. In these specifications, the size coefficients are insignificant. The effective cost coefficients remain substantially unchanged. These findings suggest that effects usually ascribed to size are more precisely attributed to trading costs. The coefficient magnitudes, however, are still troubling. The January effective cost coefficients are in the vicinity of four, implying that equilibrium gross returns reflect a turnover for the marginal investor of twice per month.

One potential explanation for the direction of the January return/liquidity seasonality might be that December closing prices are more likely to be at bid prices and January closing prices are more concentrated at asks, perhaps due to tax-related trading. The analysis of end-of-month trade directions in Section IV (and Figure 4), however, suggests otherwise, that is, that December closing prices have a slight (but statistically significant) tendency to be at the ask, while January's closing prices have little evident direction.

It might also be conjectured that the January seasonality is a spurious methodological artifact arising in some fashion from the calendar-year estimation of the effective cost and the end-of-year portfolio formation. To address this concern, I repeated all analyses with Gibbs estimations based on samples ending in June and portfolios formed at the end of June. The results are essentially unchanged.

The results for the full NYSE, Amex, and NASDAQ samples are reported in Panels B, C, and D of Table VI. The coefficients corresponding to the market prices of factor risk ( $\lambda_m$ ,  $\lambda_{smb}$ , and,  $\lambda_{hml}$ ) vary considerably across the samples and specifications. When relative size is used without effective cost, the results (not shown) uniformly confirm the strong January small-firm effect. In the specifications that include effective cost, however, the coefficient patterns are very similar to those found in the benchmark sample. In all samples, effective cost is positively (and generally significantly) associated with returns. The effect is concentrated in January, but (except for the NASDAQ sample) is also present in other months. The role of effective cost persists in the presence of (and often essentially displaces the explanatory power of) the January size variable. The point estimates of the  $c_{it}^{Gibbs}$  coefficients, however, generally imply unreasonably high equilibrium levels of turnover.

The estimates presented here use the standard CRSP monthly returns, which are usually computed from transaction prices. However, Blume and Stambaugh (1983) show that a half-spread of  $c$  causes the expected transaction-based return to overstate the true expected return by  $c^2$ . As  $c$  is usually unknown, many authors minimize the problem by eliminating securities (such as low-priced stocks) for which the problem is likely to be most severe. Here, with the present availability of  $c$  estimates it is possible to make the bias adjustment directly. The cost-adjusted excess return is defined as  $R_{it} = r_{it} - r_{ft} - (c_{it}^{Gibbs})^2$ , where  $\overline{c_{it}^{Gibbs}}$  is the portfolio average Gibbs estimate of effective cost in the preceding year. When these are used in the return regressions, estimates of the  $\overline{c_{it}^{Gibbs}}$  coefficients (and their  $t$ -statistics) are generally slightly lower, but the overall results are not materially affected.

## VI. Variation in Effective Cost

The estimation of a liquidity measure is rarely an end in itself. One usually seeks to explain liquidity variation in the cross-section (across firms) or over time, or to relate this variation to other quantities of economic interest. This section discusses several approaches to extending the basic model to accommodate liquidity covariation and risk.

To assess variation in any liquidity estimate, the simplest strategy is to partition the sample across firms and/or time periods, form estimates over the subsamples, and use the subsample values in subsequent analysis. As noted in Section I.E, however, small samples will give large weight to the prior, and in many situations the prior for effective cost is unacceptably biased. Remedies here generally follow two paths: Make the prior more informative and/or incorporate the variation directly into the model. I discuss both of these approaches below, taking as a representative situation the construction of monthly estimates of effective cost for U.S. equities, based on daily data.

### A. Informative Priors

In estimating over successive periods, last period's posterior may be very informative about current values and so may serve as a basis for the current prior. For example, one might use the mean and standard deviation of the simulated January effective cost posterior to calibrate the  $\mu$  and  $\sigma$  parameters of the half-normal density  $N^+(\mu, \sigma^2)$  used as the prior for February. Unspecified slow-moving variation in effective cost could be accommodated by inflating the  $\sigma$  parameter of the most recent posterior. It may also be useful to include recent cross-sectional data in the prior. There are many variables known to partially explain cross-sectional variation in effective cost (e.g., market capitalization, price level, volatility, and trading venue or mechanism). In a regression of posterior mean effective cost estimates against a set of these variables, the predicted values may also be useful in calibrating priors for subsequent samples.

### B. Incorporating Variation Directly into the Model

Liquidity proxies are often used in two-step analyses, where the second step (the characterization of liquidity variation) is of primary interest. Typically, for example, monthly liquidity estimates are subsequently used as regressands in specifications that allow for liquidity covariation (e.g., Pástor and Stambaugh (2003) or Acharya and Pedersen (2005)).

It is often feasible to incorporate the second-stage analysis directly into the price change specification and estimate the entire model in one step. This general technique is widely used in other financial econometrics contexts. In asset pricing applications, for example, time variation in betas and risk premia is commonly modeled by placing parametric functions, typically linear projections on conditioning variables, directly in the return specifications. (Pástor and Stambaugh (2003) use this approach for predicted liquidity betas; also see Jagannathan, Skoulakis, and Wang (2006) and references therein.)

In the present situation, this approach is simplified by the linear regression interpretation placed on equation (3). By using linear projections to model cost variation, estimation can proceed by repeated applications of the Bayesian regression model. The price change specification (3) is generalized to

$$\begin{aligned}\Delta p_{it} &= c_{it}q_{it} - c_{i,t-1}q_{i,t-1} + \beta_{im}r_{mt} + u_{it} \quad \text{for} \\ i &= 1, \dots, N \quad \text{firms and} \quad t = 2, \dots, T.\end{aligned}\quad (8)$$

Here,  $c_{it}$  denotes the cost for firm  $i$  at time  $t$ . It is convenient to assume that the  $q_{it}$  and  $u_{it}$  are independent across firms. Thus, all commonality in efficient price movements is driven by the market factor. To modify the Gibbs sampler developed for the basic market-factor model, note that at the point where we need to simulate the  $q_{it}$ , the values of  $c_{it}$  are fixed (taken as given). Thus, these draws may be accomplished with a straightforward modification of the procedure described in Section I.B: It suffices to replace all terms involving  $cq_t$  with  $c_{it}q_{it}$ .

The effective cost can be modeled as  $c_{it} = Z_{it}\gamma_i$ , where  $Z_{it}$  is a set of known conditioning variables and  $\gamma_i$  is a firm-specific coefficient vector. With this functional form, the price change may be written as

$$\Delta p_{it} = (q_{it}Z_{it} - q_{i,t-1}Z_{i,t-1})\gamma_i + \beta_{im}r_{mt} + u_{it}.\quad (9)$$

Thus, given all other variables and parameters,  $\gamma_i$  and  $\beta_i$  are regression coefficients, and their conditional draws may be obtained using the results from the Bayesian normal regression model. Since the  $u_{it}$  are assumed independent across firms, the computation may be performed separately for each firm.

### C. Latent Factors

With sufficient additional structure, it is not even essential that the conditioning variables be observable. Bayesian Gibbs sampling approaches have been applied to multivariate models involving latent factors. (Geweke and Zhou (1996) present a treatment of the APT, for example.) Common variation in effective cost can be modeled as

$$c_{it} = \gamma_{0i} + \gamma_{i1}z_t,\quad (10)$$

where  $z_t$  is an unobserved factor common to the effective costs of all firms. When this is used in (8), the expression can be reworked so that (taking the  $q_{it}$  and  $\gamma$ s as given), the  $z_t$  appear as regression coefficients. (To remove the scale indeterminacy in the product  $\gamma_{i1}z_t$  and force nonnegativity on  $c_{it}$ , it is convenient to use independent half-unit normal distributions as priors for the  $z_t$ .)

I estimate this latent common factor model for all firms in the comparison sample. The estimated  $z_t$  (based on daily data) are moderately correlated with average effective costs estimated from TAQ data. (The correlation is 0.447 at a daily frequency and 0.670 at a monthly frequency.) The quality of daily  $\gamma$

estimates is assessed by comparing them to the coefficients obtained in a one-factor analysis of TAQ estimates sampled daily. The results here are weaker. The cross-sectional (across firms) correlation between the daily-based common factor loading coefficients (the  $\gamma_{1i}$  in (10)) and the coefficients derived from TAQ is only 0.328. There is only modest improvement when the coefficients were grouped (averaged) over portfolios.

In the full 1926 to 2006 sample, when the innovations in  $z_t$  are added as explanatory variables in the return specifications (akin to (5)), the coefficients are generally insignificantly different from zero and the increase in overall explanatory power is trivial. In cross-sectional specifications, the market price of  $z_t$  risk is insignificant and of varying sign. These equivocal findings regarding the importance of effective cost variation and risk contrast with the stronger conclusions of Pástor and Stambaugh (2003), Acharya and Pedersen (2005), and Korajczyk and Sadka (2008) using different liquidity measures. Complete details are available in an earlier working paper version of this paper.

## VII. Conclusion

This paper develops a practical approach to estimating effective costs of trading using only daily data. The estimates are Bayesian and constructed using a Gibbs procedure. In a sample of firms and periods for which TAQ data are also available, the Gibbs estimates of effective cost are strongly correlated with the TAQ estimates.

The availability of estimates based on daily data makes possible analysis of the relation between trading costs and returns over a sample that is longer and more diverse than that considered in analyses that rely on high-frequency cost estimates. The analyses in this paper cover CRSP data going back to 1926. For various samples covering listed and unlisted firms over diverse time periods, the level of effective cost is found to be positively related to expected returns. However, this relation is concentrated in January. The seasonality of liquidity effects is noted in Eleswarapu and Reinganum (1993). The present analysis confirms the presence of this phenomenon in a longer and broader sample.

The January seasonality in the effective cost impact remains statistically and economically significant even in the presence of other seasonal and size-related variables. The inclusion of a simple January dummy variable raises the January returns on all stocks by the same amount. The inclusion of a January dummy interacted with a relative size measure can account for the January small-firm effect. Yet even when both types of variables are included in return specifications, the January effective cost impact remains strong, and the explanatory power of the January relative size measure usually vanishes. Taken at face value, these findings suggest that the January small-firm effect is more properly considered to be a January trading cost effect.

The magnitude of the effective cost impact, however, is too large to be consistent with straightforward trading explanations. Point estimates suggest that the gross expected monthly return impounds triple or quadruple the effective



cost. This is sensible only if the marginal agent's trading activity is much higher than implied by measured average turnover.

The approach developed in this paper can be extended to characterize variation in effective costs driven by cross-sectional and dynamic determinants. As such it is a useful tool for assessing the effects of trading costs in many situations where daily price data are available, but high-frequency data are not.

## REFERENCES

- Acharya, Viral V., and Lasse H. Pedersen, 2005, Asset pricing with liquidity risk, *Journal of Financial Economics* 77, 375–410.
- Amihud, Yakov, 2002, Illiquidity and stock returns: Cross-section and time-series effects, *Journal of Financial Markets* 5, 31–56.
- Amihud, Yakov, and Haim Mendelson, 1980, Dealership markets: Market-making with inventory, *Journal of Financial Economics* 8, 31–53.
- Amihud, Yakov, and Haim Mendelson, 1986, Asset pricing and the bid-ask spread, *Journal of Financial Economics* 17, 223–249.
- Blume, Marshall E., and Robert F. Stambaugh, 1983, Biases in computed returns, *Journal of Financial Economics* 12, 387–404.
- Brennan, Michael J., and Avanidhar Subrahmanyam, 1996, Market microstructure and asset pricing: On the compensation for illiquidity in stock returns, *Journal of Financial Economics* 41, 441–464.
- Campbell, John Y., Sanford J. Grossman, and Jiang Wang, 1993, Trading volume and serial correlation in stock returns, *Quarterly Journal of Economics* 108, 905–939.
- Carlin, Bradley P., and Thomas A. Louis, 2000, *Bayes and Empirical Bayes Methods for Data Analysis* (Chapman and Hall, London).
- Chalmers, John M. R., and Gregory B. Kadlec, 1998, An empirical examination of the amortized spread, *Journal of Financial Economics* 48, 159–188.
- Cochrane, John H., 2005, *Asset Pricing* (Princeton University Press, Princeton).
- Easley, David, Soeren Hvidkjaer, and Maureen O'Hara, 2002, Is information risk a determinant of asset returns?, *Journal of Finance* 57, 2185–2221.
- Easley, David, and Maureen O'Hara, 2002, Microstructure and asset pricing, in George M. Constantinides, Milton Harris, and Rene M. Stulz, eds.: *Handbook of Financial Economics* (Elsevier, New York).
- Eleswarapu, Venkat R., and Marc R. Reinganum, 1993, The seasonal behavior of the liquidity premium in asset pricing, *Journal of Financial Economics* 34, 373–386.
- Elton, Edwin J., Martin J. Gruber, Joel Krasny, and Sadi Ozelge, 2006, The effect of the frequency of holding data on conclusions about mutual fund behavior, Working paper, Stern School, New York University.
- Fama, Eugene F., and Kenneth R. French, 2004, New lists: Fundamentals and survival rates, *Journal of Financial Economics* 72, 229–269.
- Fink, Jason, Kristin Fink, Gustavo Grullon, and James P. Weston, 2006, Firm age and fluctuations in idiosyncratic risk, Working paper, Jones School, Rice University.
- Geweke, John, 2005, *Contemporary Bayesian Statistics and Econometrics* (John Wiley and Sons, New York).
- Geweke, John, and Guofu Zhou, 1996, Measuring the pricing error of the arbitrage pricing theory, *Review of Financial Studies* 9, 557–587.
- Goyenko, Ruslan, Craig W. Holden, Christian T. Lundblad, and Charles A. Trzcinka, 2005, Horse-races of monthly and annual liquidity measures, Working paper, Indiana University.
- Hajivassiliou, Vassilis, Daniel McFadden, and Paul Ruud, 1996, Simulation of multivariate normal rectangle probabilities and their derivatives—Theoretical and computational results, *Journal of Econometrics* 72, 85–134.

- Harris, Lawrence E., 1990, Statistical properties of the Roll serial covariance bid/ask spread estimator, *Journal of Finance* 45, 579–590.
- Hasbrouck, Joel, 2004, Liquidity in the futures pits: Inferring market dynamics from incomplete data, *Journal of Financial and Quantitative Analysis* 39, 305–326.
- Hasbrouck, Joel, and George Sofianos, 1993, The trades of market makers: An empirical examination of NYSE specialists, *Journal of Finance* 48, 1565–1593.
- He, Jia, Lilian Ng, and Qinghai Wang, 2004, Quarterly trading patterns of financial institutions, *Journal of Business* 77, 493–510.
- Jagannathan, Ravi, Georgios Skoulakis, and Zhenyu Wang, 2006, The analysis of the cross-section of security returns, in Lars Hansen and Yacine Aït-Sahalia, eds.: *Handbook of Financial Econometrics* (Elsevier North-Holland).
- Korajczyk, Robert A., and Ronnie Sadka, 2008, Pricing the commonality across alternative measures of liquidity, *Journal of Financial Economics* 87, 45–72.
- Lakonishok, Josef, Richard H. Thaler, and Robert Vishny, 1991, Window dressing by pension fund managers, *American Economic Review* 81, 227–231.
- Lancaster, Tony, 2004, *An Introduction to Modern Bayesian Econometrics* (Blackwell Publishing, Malden, MA).
- Lee, Charles M. C., 1993, Market integration and price execution for NYSE-listed securities, *Journal of Finance* 48, 1009–1038.
- Lesmond, David A., Joseph P. Ogden, and Charles A. Trzcinka, 1999, A new estimate of transactions costs, *Review of Financial Studies* 12, 1113–1141.
- Lo, Andrew W., Harry Mamaysky, and Jiang Wang, 2004, Asset prices and trading volume under fixed transaction costs, *Journal of Political Economy* 112, 1054–1090.
- Madhavan, Ananth, and Seymour Smidt, 1993, An analysis of changes in specialist inventories and quotations, *Journal of Finance* 48, 1595–1628.
- Meier, Iwan, and Ernst Schaumburg, 2004, Do funds window dress? Evidence for U.S. equity mutual funds, Working paper, Northwestern University.
- Musto, David, 1997, Portfolio disclosures and year-end price shifts, *Journal of Finance* 52, 1563–1588.
- Musto, David, 1999, Investment decisions depend on portfolio disclosures, *Journal of Finance* 54, 935–952.
- O'Neal, Edward, 2001, Window dressing and equity mutual funds, Working paper, Wake Forest University.
- Pástor, Ľuboš, and Robert F. Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642–685.
- Piowar, Michael S., and Li Wei, 2006, The sensitivity of effective spread estimates to trade-quote matching algorithms, *International Journal of Electronic Markets* 16, 112–129.
- Reinganum, Marc R., 1990, Market microstructure and asset pricing: An empirical investigation of NYSE and NASDAQ securities, *Journal of Financial Economics* 28, 127–147.
- Roll, Richard, 1984, A simple implicit measure of the effective bid-ask spread in an efficient market, *Journal of Finance* 39, 1127–1139.
- Sadka, Ronnie, 2004, Liquidity risk and asset pricing, Working paper, University of Washington.
- Schultz, Paul H., 2000, Regulatory and legal pressures and the costs of Nasdaq trading, *Review of Financial Studies* 13, 917–957.
- Shanken, Jay, 1992, On the estimation of beta pricing models, *Review of Financial Studies* 5, 1–33.
- Sias, Richard, 2006, Window dressing, tax-loss selling and momentum profit seasonality, Working paper, Washington State University.
- Sias, Richard, and Laura Starks, 1997, Institutions and individuals, *Journal of Finance* 52, 1543–1562.
- Spiegel, Matthew, and Xiatong Wang, 2005, Cross-sectional variation in stock returns: Liquidity and idiosyncratic risk, Working paper, Yale University.
- Stoll, Hans R., 2006, Electronic trading in stock markets, *Journal of Economic Perspectives* 20, 153–174.
- Stoll, Hans R., and Robert E. Whalley, 1983, Transaction cost and the small firm effect, *Journal of Financial Economics* 12, 57–79.