ESSAYS ON LIQUIDITY, LIQUIDITY RISK, AND METHOD OF MOMENTS ESTIMATION

by

Roni Israelov

Submitted in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy in the

Department of Finance

Tepper School of Business
Carnegie Mellon University
May 2007

Dissertation Committee:
Professor Burton Hollifield (chair)
Professor Carolyn Levine
Professor Duane Seppi
Professor Fallaw Sowell
Acknowledgments

This dissertation would not exist had it not been for the advice and generosity of Dr. Sanjay Srivastava. The doctoral path is one that I had not seriously considered until our discussions and without his encouragement I would not have embarked on the journey that has encompassed the past four years of my life. Dr. Srivastava paved the path for me to Carnegie Mellon University and ensured the availability of funding for my graduate program. For this, I will be forever grateful.

When I think back on all the meetings and conversations I have had with my advisor, Dr. Burton Hollifield, I am just stunned. Week after week, month after month, and year after year, Dr. Hollifield encouraged me to follow my own interests, listened to my ideas, critiqued my work, and provided me with direction when I was stuck. One cannot ask for a more genuinely helpful advisor and I cannot even begin to imagine what the past four years would have been like without him.

Four years ago, I thought it impossible that econometrics could interest me. Yet, somehow, through Dr. Fallaw Sowell, I became interested enough to devote one chapter of my dissertation to the topic. This feat required the help of an extremely gifted individual whose passion for econometrics is contagious. Dr. Sowell is one of the best teachers that I have had in my life and I am grateful for his advice, encouragement, and willingness to help.

I am grateful to Dr. Duane Seppi, whose comments and suggestions have had tremendous influence on the first two chapters of my dissertation. I gained a great deal of insight into my own work through his questions and my papers would not be complete with his guidance.

As a graduate student working on his dissertation, it is easy to lose sight of the forest for the trees. I am grateful for Dr. Richard Green and his vast understanding of the field of finance. I know our conversations helped me develop the proper perspective for my work.

I would also like to thank Dr. Michael Gallmeyer, Dr. Shimon Kogan, Dr. Bryan Routledge, and Dr. Chris Telmer for numerous discussions and advice over the past few years.

The life of a graduate student is an emotional roller coaster that results in a very special bond between classmates. Without naming names, you know who you are, and I am honored to have traveled this path alongside you.

To my wife: You had absolutely no idea what lay ahead when I asked what you thought about moving to Pittsburgh so that I may pursue a Ph.D. at Carnegie Mellon University. Yet, without hesitation, you offered your full and unconditional support. It has been a wild ride. Thank you!
Abstract

Because prices are a function of expected future liquidity levels, a component of unexpected returns is due to new information about future costs to trade. The first two chapters of this dissertation explore this relationship.

Chapter 1 extends the Campbell (1991) return decomposition to include liquidity news, which captures the component of unexpected returns due to revisions in future expected liquidity levels. For the market portfolio of NYSE and AMEX stocks between 1964 and 2001, liquidity news has approximately 5 times the volatility of contemporaneous liquidity and 1.2 percent the volatility of contemporaneous returns. Liquidity news’s contribution to portfolio volatility is higher for small and illiquid stocks. Systematic liquidity news risk is priced and has explanatory power for the cross-sectional variation in expected returns. However, the price of systematic liquidity news risk is not statistically different than that of dividend and discount rate news risk.

Chapter 2 investigates the cross-sectional variation in changes in valuation after the Tokyo Stock Exchange shut down on January 18, 2006 because its daily capacity of 4.5 million trades had been reached. According to simulations using historical data, 1.5 additional closures are expected over the six month period following the event. Stocks traded only on the TSE lost 2 percent of their value relative to those traded on at least one additional Japanese exchange and liquid stocks lost value relative to illiquid stocks. For instance, high turnover stocks lost approximately 6 percent of their value relative to low turnover stocks.

Under small samples, numerous or unbounded moments, or mis-specified models, Empirical Likelihood (EL)’s estimator may have restricted support and may be undefined at or near the population parameter value. Chapter 3 proposes and investigates the properties of an estimator named Penalized Method of Moments (PMM) that merges the GMM and EL objective functions. Through a free parameter, PMM’s estimator properties, including its support, may be adjusted with GMM and EL behavior attained at the extremes. For a simulated Schennach (2006) model with small samples and numerous moments, PMM’s estimates are less volatile and its hypothesis tests are less mis-sized than those of EL.
# Table of Contents

Acknowledgments ii

Abstract iii

1 Future Liquidity, Present Value: Measuring and Pricing Liquidity Risk 2
   1.1 Introduction .......................................................... 2
   1.2 Decomposing Asset Returns ........................................ 6
       1.2.1 Fixed Cost Decomposition .................................. 7
       1.2.2 Proportional Cost Decomposition ......................... 9
       1.2.3 Decomposing the Net Return Variance ..................... 13
   1.3 Vector Autoregressions ............................................ 14
   1.4 State Variables, Data Set, and Illiquidity Proxy ............ 17
       1.4.1 A Measure of Trading Costs ............................ 19
       1.4.2 Inclusion Requirements .................................. 20
       1.4.3 Portfolio Formation ....................................... 21
       1.4.4 Calculating the News Coefficients ....................... 22
   1.5 Results of the VAR Decompositions ............................ 23
       1.5.1 Proportional Cost Decomposition ....................... 24
       1.5.2 Fixed Cost Decomposition ............................... 25
       1.5.3 Cross-Sectional Variation in the Variance Decompositions .... 26
       1.5.4 Regime Shifts and Parameter Stability .................. 28
   1.6 Pricing Liquidity Risk ........................................... 30
1.6.1 Estimating the Betas ................................................. 32
1.6.2 The Liquidity Risk–Return Relationship ......................... 35
1.6.3 Economic Significance .............................................. 38
1.7 Conclusion ............................................................. 40

2 Surprise! It’s Illiquid: ................................................. 55
2.1 Introduction ........................................................... 55
2.2 Background Information ............................................. 59
  2.2.1 The Exchanges .................................................... 59
  2.2.2 The Liquidity Event .............................................. 61
2.3 Empirical Analysis .................................................... 62
  2.3.1 Data ............................................................... 62
  2.3.2 Likelihood of Future TSE Shutdown ............................ 62
  2.3.3 Relative Value Changes ......................................... 64
  2.3.4 Measured Liquidity Shifts ....................................... 74
2.4 Conclusion ............................................................. 76

3 Penalized Method of Moments and Empirical Likelihood Estimation: 95
3.1 Introduction ........................................................... 95
3.2 Definitions and Notation ............................................. 98
3.3 The Ingredients: GMM and EL ....................................... 99
  3.3.1 Generalized Method of Moments ............................... 99
  3.3.2 Empirical Likelihood ............................................ 101
3.4 Penalized Method of Moments ..................................... 104
3.5 EL, Negative Weights, and Shrinkage ............................. 111
3.6 Computation .......................................................... 114
3.7 Monte Carlo Simulations ............................................ 115
3.8 Conclusion ............................................................. 117

Bibliography ............................................................ 131

Appendix ................................................................. 137
Chapter 1

Future Liquidity, Present Value: Measuring and Pricing Liquidity Risk

1.1 Introduction

Amihud, Mendelson, and Pedersen (2005) write in their extensive survey of the liquidity literature:

“Liquidity varies over time. This means that investors are uncertain what transactions cost they will incur in the future when they need to sell an asset. Further, since liquidity affects the level of prices, liquidity fluctuations can affect the asset price volatility itself.”

This paper joins the growing body of research that studies the effect of liquidity risk on expected returns, but it begins with an investigation of the effect of liquidity fluctuations on asset price volatility. Amihud (2002) reports that negative shocks to liquidity lower asset prices. Acharya and Pedersen (2005) report a negative relationship between unexpected market illiquidity and asset returns, and between unexpected asset illiquidity and market returns. How much of an asset’s volatility can be attributed to these liquidity fluctuations?

A similar question can be and has been asked about dividend fluctuations. Campbell (1991) answers the question by decomposing the unexpected contemporaneous return into two
components, new information about the dividend stream and new information about future discount rates. Even though uncertainty about contemporaneous cash flows is relatively low, using the decomposition, Campbell (1991) reports that new information about future cash flows accounts for almost half the variation of contemporaneous returns. I extend the Campbell (1991) return decomposition to include a liquidity component in two ways. The first decomposition considers revisions in the expected growth of fixed costs and the second decomposition considers new information about proportional costs. Each measure captures the resolved uncertainty about contemporaneous liquidity costs as well as the price impact of new information about future fixed and proportional costs.

I estimate the two decompositions for an equal-weight aggregate portfolio of NYSE and AMEX stocks over the period January 1964 to December 2001 using the proportional cost proxy implemented by Acharya and Pedersen (2005), which is a variation of the Amihud (2002) illiquidity measure. I find that the price impact of new information about future liquidity is a larger source of portfolio risk than uncertainty about contemporaneous proportional costs. Although proportional cost news has roughly 5.2 times the volatility of contemporaneous proportional costs, its volatility is about 100 times smaller than that of contemporaneous returns. Hence, liquidity risk, from the proportional cost perspective, does not appear to be an economically significant source of risk for the aggregate portfolio. Fixed cost news, on the other hand, is approximately 105 times more volatile than contemporaneous proportional costs and has 25 percent the volatility of contemporaneous returns. I derive an explicit relationship between news about fixed costs, proportional costs, and dividends. The relationship suggests that fixed cost news is primarily driven by new information about future dividends. Positive dividend news leads to a capital gain – if proportional costs do not adjust, the increased price results in higher fixed costs.

In addition to the aggregate portfolio, I study the properties of fixed and proportional cost news over the cross-section by forming quintile-ranked portfolios after sorting assets on market capitalization, illiquidity levels, and turnover. The liquidity news components of these portfolios’ returns have similar properties to those of the aggregate portfolio. Proportional cost and fixed cost news are significantly more volatile than contemporaneous proportional
costs and fixed cost news is substantially more volatile than proportional cost news. I report that small and illiquid stocks have more volatile proportional and fixed cost news and low turnover stocks have more volatile proportional cost news and less volatile fixed cost news.

There is growing evidence that systematic liquidity risk is priced. Pastor and Stambaugh (2003) report that assets with returns that are sensitive to market liquidity earn a risk premium. Sorting assets by their return sensitivities to the aggregate liquidity measure, the portfolio that is long stocks in the highest decile and short stocks in the lowest decile has annualized returns of 8.5 percent after adjusting for the Fama-French factors. Acharya and Pedersen (2005) derive a Liquidity-Adjusted CAPM (LACAPM) that includes a market beta and three liquidity betas. They report that the total liquidity risk premium is approximately 1.1 percent when all sources of portfolio systematic risk are restricted to have the same market price of risk. Allowing the prices of the three liquidity risks to differ from gross return risk, depending on the specification, Acharya and Pedersen (2005) report the estimated price of liquidity risk to be between 4 and 22 times larger than that of gross return risk. Sadka (2006) reports a positive and statistically significant liquidity risk premium using the Fama-French model augmented by a liquidity factor. His liquidity factor is the average permanent market impact coefficient after decomposing a measure of Kyle’s $\lambda$ into permanent and transitory effects. The sensitivity Sadka (2006) measures is similar to that measured by Pastor and Stambaugh (2003) and is related to one of the three liquidity betas in the Acharya and Pedersen (2005) model.

The second half of the paper explores the relationship between expected returns and liquidity news risk by applying the two net return decompositions to the Acharya and Pedersen (2005) LACAPM. A simple description of LACAPM is that it is CAPM, except the gross return is replaced by the net return. An asset’s expected excess net return is its net return beta times the market expected excess net return. The final version of their model is obtained by substituting in the relationship that net return equals gross return minus proportional costs. In their estimation, returns are assumed to be independent over time and proportional costs are assumed to be a martingale. Relaxing these assumptions by accounting for return momentum and liquidity mean reversion, I report (i) a slight drop
in the explanatory power of the LACAPM and (ii) the market price of liquidity risk is no longer statistically different than non-liquidity risk.

In addition to their specification, I re-estimate the LACAPM under the two alternative specifications obtained through the proportional and fixed cost decompositions. The Acharya and Pedersen (2005) specification restricts the market price of the liquidity risk associated with new information about future liquidity to equal that of dividend news and discount rate news. The alternate specifications obtained through the net return decomposition restrict the market price of the two sources of liquidity risk to the same value.

I find that systematic liquidity news risk, as defined by the two decompositions and estimated through the LACAPM, helps to explain the cross-sectional variation in expected stock returns. I estimate the liquidity premium between the least liquid and most liquid stocks to be approximately 4.1 percent per year. Systematic proportional cost news risk accounts for an additional 0.6 percent per year difference across the two groups of stocks. The total premium associated with liquidity – level and risk – is approximately 4.7 percent per year.

The rest of the paper is as follows. The next section provides the decomposition of asset net returns into their three components. Section 3 sets up the vector autoregression and relates the VAR to the variance decomposition provided in section 2. In section 4, I describe the data, the data inclusion requirements, and the construction of the cross-sectional portfolios analyzed in the paper. Section 5 reports the results of the variance decomposition for the aggregate and characteristic-sorted portfolios. Section 6 estimates the price of liquidity news risk. Section 7 concludes and the Appendix includes details on the jackknife resampling technique used in this paper.
1.2 Decomposing Asset Returns

My decompositions begin with the Campbell (1991) log-linear approximation of unexpected returns:

\[ r_t - E_{t-1}r_t \approx \Delta E_t \sum_{i=0}^{\infty} \rho^i \Delta c f_{t+i} - \Delta E_t \sum_{i=1}^{\infty} \rho^i r_{t+i}, \quad (1.1) \]

where \( r_t \) is the log stock return, \( c f_t \) is the log cash flow paid by the stock, \( \Delta \) denotes a one-period change, \( E_t \) denotes a rational expectation at time \( t \), and \( \rho \equiv P/(P + D) \) is a discount coefficient defined at long-horizon means. Equation (1.1) is the result of a first-order approximation of an accounting identity and not an economic or behavioral model. It says that a capital gain today requires that expected future cash flow growth be higher, or expected future asset returns be lower, or both.

In the Campbell (1991) model, a stock’s cash flows are its dividend stream. The notation is simplified by defining dividend news \((\eta_{d,t})\), discount rate news \((\eta_{r,t})\), and unexpected returns \((\nu_t)\) as follows:

\[ \eta_{d,t} \equiv \Delta E_t \sum_{i=0}^{\infty} \rho^i d_{t+i} \quad \eta_{r,t} \equiv \Delta E_t \sum_{i=1}^{\infty} \rho^i r_{t+i} \quad \nu_t \equiv r_t - E_{t-1}r_t. \quad (1.2) \]

Then (1.1) can be rewritten in compact form as \( \nu_t \approx \eta_d - \eta_r \).

I introduce illiquidity by considering the net or spread-adjusted return in a manner similar to Amihud and Mendelson (1986), Jones (2002), and Acharya and Pedersen (2005):

\[ \tilde{R}_t \equiv \frac{P_t + D_t - C_t}{P_{t-1}} = R_t - K_t, \quad (1.3) \]

where \( C_t \) is the contemporaneous per share fixed cost incurred each period for holding the illiquid asset and \( K_t \equiv C_t/P_{t-1} \) is the contemporaneous proportional cost. Amihud, Mendelson, and Pedersen (2005) describe the following four sources of illiquidity: exogenous transactions costs (commissions, taxes, etc.), demand pressure and inventory risk\(^1\),

asymmetric information about asset fundamentals or order flow\(^2\), and search frictions\(^3\). I assume the fixed cost is exogenous and captures the four illiquidity sources. I also assume the cost is incurred each period regardless of whether a transaction takes place. The per period cost includes the monetized opportunity cost associated with holding the illiquid asset, which may be due to second-best portfolio optimization or the inability to profit from small informational asymmetries. It also amortizes over the holding period the illiquidity cost incurred when the asset is liquidated.

### 1.2.1 Fixed Cost Decomposition

To the marginal investor, the realized cash flow at time \( t \) is the dividend paid by the stock minus liquidity costs. The log cash flow may be written \( cf_t \equiv \log(e^{dt} - e^{ct}) \). Note that log dividends \( (d_t) \) and log costs \( (c_t) \) enter \((1.1)\) nonlinearly through \( cf_t \). A first-order Taylor expansion around long run means gives the following approximation of cash flow growth rates

\[
\Delta cf_t \approx \omega_d \Delta d_t - \omega_c \Delta c_t,
\]

where \( \omega_d \) and \( \omega_c \) are defined to be long run means of the following ratios:

\[
\omega_d = \frac{D}{D-C}, \quad \omega_c = \frac{C}{D-C}.
\]

We may think of an asset as a portfolio that is long \( \omega_d \) shares of a dividend asset and short \( \omega_c \) shares of an illiquidity cost asset. Then the cash flow growth of the “portfolio” is the weighted average growth of its two holdings, dividends and costs.


Substituting (1.4) into (1.1), I obtain

\[ \tilde{\tau}_t - E_{t-1} \tilde{\tau}_t \approx \omega_d \Delta E_t \sum_{i=0}^{\infty} \rho^i \Delta d_{t+i} - \omega_c \Delta E_t \sum_{i=0}^{\infty} \rho^i \Delta c_{t+i} - \Delta E_t \sum_{i=1}^{\infty} \rho^i \tilde{\tau}_{t+i}, \]  

(1.6)

which may be written in compact form to simplify notation

\[ \tilde{\nu}_t \approx \omega_d \eta_{d,t} - \omega_c \eta_{c,t} - \eta_{\tilde{\tau},t} \]  

(1.7)

\[ = \hat{\eta}_{d,t} - \hat{\eta}_{c,t} - \hat{\eta}_{\tilde{\tau},t}, \]  

(1.8)

where \( \hat{\eta}_{d,t} \equiv \omega_d \eta_{d,t}, \hat{\eta}_{c,t} \equiv \omega_c \eta_{c,t} \), and fixed cost news is defined by

\[ \eta_{c,t} = \Delta E_t \sum_{i=0}^{\infty} \rho^i \Delta c_{t+i}. \]  

(1.9)

Equations (1.6) through (1.8) relate the unexpected component of net returns to revisions in expectations about future dividend growth, fixed cost growth, and net discount rates. New information about higher than expected future fixed costs negatively influences contemporaneous prices. News about dividends and discount rates have the same effect on unexpected returns as in the Campbell (1991) model. Dividend and fixed cost news are multiplied by \( \omega_d \) and \( \omega_c \) respectively in order to account for their relative importance on prices. The natural assumption that long-run dividends are greater than long-run transactions costs dictates that the coefficient on dividend news is greater than that on fixed cost news. News that future dividends will increase by ten percent is more important than news that future costs will increase by ten percent. Both coefficients increase as the long-run mean of the cost-dividend ratio increases and achieve their minimum values when trading costs are absent. Cross-sectionally, these coefficients are important; for example, small firms tend to have low yields and high trading costs and large firms have high yields and low trading costs.

---

4An exact substitution of the log cash flow approximation into the Campbell (1991) decomposition would lead to \( \rho = \frac{\hat{\rho}}{P / D - c} \). Expanding the price-cash flow ratio in the Campbell and Shiller (1988) approximation around the long run mean price-dividend ratio instead of the price-cash flow ratio leads to \( \rho = \frac{\rho}{P + D} \) as defined in this paper.
1.2.2 Proportional Cost Decomposition

An alternate decomposition considers revisions in expected proportional rather than fixed costs. The log gross return of a stock may be approximated by $r_t \approx \tilde{r}_t + K_t$. Substituting the log gross return approximation into the Campbell (1991) decomposition provided by equation (1.1), I obtain

$$\tilde{r}_t - E_{t-1} \tilde{r}_t \approx \Delta E_t \sum_{i=0}^{\infty} \rho^i \Delta d_{t+i} - \Delta E_t \sum_{i=0}^{\infty} \rho^i K_{t+i} - \Delta E_t \sum_{i=0}^{\infty} \rho^i \tilde{r}_{t+i}$$

(1.10)

which may be written in compact form to simplify notation

$$\tilde{r}_t \approx \eta d, t - \eta K, t - \eta \tilde{r}, t,$$

(1.11)

where proportional cost news, $\eta K, t$ is defined as follows:

$$\eta K, t = \Delta E_t \sum_{i=0}^{\infty} \rho^i K_{t+i} = \Delta E_t K_t + \sum_{i=1}^{\infty} \rho^i K_{t+i}.$$ 

(1.12)

In the proportional cost decomposition, cash flows are the dividend stream so that transactions costs are not double counted. Equations (1.10) and (1.11) relate the unexpected net return to changes in expected dividend growth, proportional costs, and net returns. Negative contemporaneous returns occur when proportional cost estimates are revised upwards. From equation (1.12), we see that proportional cost news includes the unexpected contemporaneous proportional cost, which is the source of liquidity risk identified by Acharya and Pedersen (2005). In addition, proportional cost news also contains the newly acquired information about future proportional costs. This additional source of liquidity risk, the price impact of new information about future liquidity, is due to the time-series variation in expected liquidity.

Fixed costs and proportional costs are obviously related. Equations (1.8) and (1.11) may be combined to relate the two liquidity news terms. Two equivalent forms of the relationship
follow

\[ \eta_{c,t} = \eta_{d,t} + \ddot{\eta}_{K,t} \quad (1.13) \]
\[ \eta_{K,t} = \omega_c (\eta_{c,t} - \eta_{d,t}) \quad (1.14) \]

where \( \ddot{\eta}_{K,t} \equiv \eta_{K,t}/\omega_c \). Proportional costs link information about dividends to information about fixed costs. If proportional costs are time-invariant or there is no newly acquired information about proportional costs, then \( \ddot{\eta}_{K,t} = 0 \) and \( \eta_{c,t} = \eta_{d,t} \). With no change in proportional costs, a 10 percent increase in dividends implies a 10 percent increase in fixed costs. Proportional cost news is primarily driven by price changes. When expected future fixed costs are higher \( (\eta_{c,t} > 0) \) or dividends are lower \( (\eta_{d,t} < 0) \), prices drop and expected proportional costs increase \( (\eta_{K,t} > 0) \). A similar equation may be derived for yield news:

\[ \eta_{Y,t} \equiv \Delta E_t \sum_{i=0}^{\infty} \rho^i Y_{t+i} = \omega_d (\eta_{c,t} - \eta_{d,t}). \quad (1.15) \]

The contemporaneous yield is defined to be the contemporaneous dividend divided by lag price: \( Y_t \equiv D_t/P_{t-1} \). As is the case for proportional cost news, yield news is also primarily driven by price changes. Positive fixed cost news or negative dividend news lowers price and increases the yield. Yield and proportional cost news are related as follows:

\[ \eta_{Y,t} = \frac{\omega_d}{\omega_c} \eta_{K,t} = \frac{D}{C} \eta_{K,t}. \quad (1.16) \]

Equation (1.16) predicts that the dividend yield should be more volatile than the proportional cost, a prediction that is consistent with the empirical findings reported by Jones (2002). Using annual data for the entire twentieth century, Jones (2002) regresses yield on proportional costs and reports that a one unit increase in proportional costs is associated with a 5.3 unit increase in yield. Although the regression does not translate perfectly to the relationship given by equation (1.16), the estimate suggests that dividends are approximately 5.3 times larger than costs. Chalmers and Kadlec (1998) report that between 1983 and 1992 the equal-weight amortized spread was approximately 0.51 percent per year. The equal-weight yield over the same period was approximately 2.82 percent per year. Hence,
the dividend-cost ratio during the ten year period was approximately 5.5. Together, the findings suggest that fixed costs are roughly 18 percent the size of dividends. Ceteris paribus, new information about fixed costs should have $\frac{18}{1.8} = 15.3$ percent the effect on returns as new information about dividends.

Although the three-term decompositions distinguish the three sources of return information and risk, the primary purpose of this paper is to investigate the properties of the liquidity component. In Section 6, I estimate the market prices of fixed cost versus non-fixed cost news risk and of proportional cost versus non-proportional cost news risk. In order to do so, I aggregate the non-liquidity components in each decomposition to provide two-term decompositions that simplify the analysis and exposition considerably. Defining $\eta^*_{c,t} \equiv \tilde{\eta}_{d,t} - \tilde{\eta}_{F,t}$ to be non-fixed cost news and $\eta^*_{K,t} \equiv \eta_{d,t} - \eta_{F,t}$ to be non-proportional cost news, the unexpected component of net returns may be rewritten $\tilde{\nu}_t \approx \eta^*_{c,t} - \eta_{c,t}$ and $\tilde{\nu}_t \approx \eta^*_{K,t} - \eta_{K,t}$.

**Proportional Cost News: Two Examples**

The remainder of the paper explores the contribution of information about liquidity on portfolio volatility during the last forty years of the previous century. The decompositions presented above allow for more than data description. They tell us how new information about liquidity (and dividends and discount rates) are incorporated into prices. I now consider the unexpected returns associated with two recent examples of persistent liquidity shocks.

In 1971, the NYSE began the process of commission deregulation on orders above $500,000. The process continued and extended to smaller orders until May 1, 1975 when all commissions were deregulated. Even before deregulation began, a small volume discount was provided on transactions over 1000 shares beginning in December 1968. Jones (2002) reports the average commissions for NYSE stocks beginning in the mid 1920s until the end of the century. The average one-way commission between 1960 and 1976 was approximately 80 basis points. Jones (2002) reports that beginning in 1976, commissions rapidly dropped with a half-life of approximately 7.5 years. Suppose investors correctly predicted the change
in trading costs associated with commission deregulation. What is the change in market valuation associated with the information?

The example under consideration is not a single persistent shock to trading costs. Commissions continue a steady decline over time. In order to model the trading cost, I assume that prior to deregulation future expected round-trip trading costs are constant and amortized to be $K_0 = 0.0768$ percent per month, consistent with a one-way commission of 80 basis points and holding period of approximately 21 months. After deregulation, trading costs follow a mean-reverting process with long-horizon mean $\bar{K} = 0.0096$ percent per month and persistence $\delta = 0.992328$, which corresponds to a half-life of approximately 7.5 years. The new information is due to the shift in behavior for trading costs. After a few steps of algebra, it can be shown that proportional cost news at the time of information revelation is:

$$\eta_{K,0} = -\left(\frac{\rho}{1-\rho\delta}\right)\left(\frac{1-\delta}{1-\rho}\right)(K_0 - \bar{K}).$$ \hfill (1.17)

Substituting in the values provided in the above paragraph and $\rho = 0.95^{\frac{1}{12}}$, I calculate the capital gain associated with the degradation in commissions to be approximately 10 percent.

A second example is the decimalization of prices on the NYSE on January 29, 2001 and on the NASDAQ on April 9, 2001. The intuition that the reduction in tick size would be associated with a narrowing of bid-ask spreads is confirmed by a number of articles investigating the decimalization of the Canadian stock exchanges in April 1996; see Ahn, Cao, and Choe (1997), Bacidore (1997), Huson, Kim, and Mehrotra (1997), and Ricker (1997) who all document an almost instantaneous reduction in spreads following decimalization. For the decimalization of the NYSE, Bessembinder (2002) reports an immediate drop in effective spreads from 0.92 to 0.52 percent. Converting the spread to a monthly amortized cost provides a drop of 0.0192 percent per month. According to the proportional cost decomposition and the definition of proportional cost news, the information associated with the spread reduction is associated with a capital gain of $0.0192/(1-\rho) = 4.5$ percent.
1.2.3 Decomposing the Net Return Variance

The three-term return decompositions result in six-term variance decompositions:

\[
\text{var}(\tilde{\nu}_t) = \text{var}(\tilde{\eta}_{dt,t}) + \text{var}(\tilde{\eta}_{ct,t}) + \text{var}(\eta_{\tilde{r},t}) - 2 \text{cov}(\tilde{\eta}_{dt,t}, \tilde{\eta}_{ct,t}) - 2 \text{cov}(\tilde{\eta}_{dt,t}, \eta_{\tilde{r},t}) + 2 \text{cov}(\tilde{\eta}_{ct,t}, \eta_{\tilde{r},t})
\]

(1.18)

for the fixed cost decomposition and

\[
\text{var}(\tilde{\nu}_t) = \text{var}(\eta_{dt,t}) + \text{var}(\eta_{K,t}) + \text{var}(\eta_{\tilde{r},t}) - 2 \text{cov}(\eta_{dt,t}, \eta_{K,t}) - 2 \text{cov}(\eta_{dt,t}, \eta_{\tilde{r},t}) + 2 \text{cov}(\eta_{K,t}, \eta_{\tilde{r},t})
\]

(1.19)

for the proportional cost decomposition. An objective of the above variance decompositions is to determine liquidity’s influence on portfolio volatility. The task is ambiguous because liquidity enters the above variance decompositions in three places, through one variance term and two covariance terms. There is a continuum of measures that aggregate the variance and covariance terms to capture different aspects of liquidity’s influence on return volatility. I report the following three measures for each decomposition. The first measure is the ratio of cost volatility to return volatility: \( \tilde{R}_c \equiv \sigma(\tilde{\eta}_{ct,t})/\sigma(\tilde{\nu}_t) \) for fixed costs and \( \tilde{R}_K \equiv \sigma(\eta_{K,t})/\sigma(\tilde{\nu}_t) \) for proportional costs. The measure is a relative measure of liquidity news volatility and does not take into account any relationship between liquidity news and the other two news terms. The relative volatilities of dividend news and discount rate news are similarly defined by \( \tilde{R}_d \equiv \sigma(\eta_{dt,t})/\sigma(\tilde{\nu}_t) \) for dividend news in the fixed cost decomposition, \( \tilde{R}_d \equiv \sigma(\eta_{dt,t})/\sigma(\tilde{\nu}_t) \) for dividend news in the proportional cost decomposition, and \( \tilde{R}_\tilde{r} \equiv \sigma(\eta_{\tilde{r},t})/\sigma(\tilde{\nu}_t) \) for net discount rate news. The second measure I report is the regression coefficient of liquidity news on unexpected contemporaneous returns: \( \tilde{B}_c \equiv \text{cov}(\tilde{\nu}_t, \tilde{\eta}_{ct,t})/\text{var}(\tilde{\nu}_t) \) for fixed costs and \( \tilde{B}_K \equiv \text{cov}(\tilde{\nu}_t, \eta_{K,t})/\text{var}(\tilde{\nu}_t) \) for proportional costs. When the measure is positive, unexpected contemporaneous returns are less volatile than non-liquidity news because of the positive comovement between liquidity and non-liquidity news. Similarly, I calculate the regression coefficients for dividend and net return news: \( \tilde{B}_d \equiv \text{cov}(\tilde{\nu}_t, \eta_{dt,t})/\text{var}(\tilde{\nu}_t) \), \( \tilde{B}_d \equiv \text{cov}(\tilde{\nu}_t, \eta_{dt,t})/\text{var}(\tilde{\nu}_t) \) and \( \tilde{B}_\tilde{r} \equiv \text{cov}(\tilde{\nu}_t, \eta_{\tilde{r},t})/\text{var}(\tilde{\nu}_t) \).
By definition, $B_d - B_c - B_r = 1$ and $B_d - B_K - B_r = 1$. These regression coefficients are similarly defined as the measures of dividend and discount rate contribution to portfolio variance reported by Campbell (1991) and can be loosely interpreted as the percentage of conditional return variance attributable to the respective return component. The final reported measure is the ratio of cost volatility to contemporaneous proportional cost volatility: $\hat{P}_c \equiv \sigma(\bar{\eta}_{c,t})/\sigma(K_t)$ for proportional costs and $P_K \equiv \sigma(\bar{\eta}_{K,t})/\sigma(K_t)$. The measure captures how large the volatility of liquidity news is relative to that of uncertainty about contemporaneous proportional costs, the traditional measure of liquidity risk and may be interpreted as a measure of liquidity’s persistence. Campbell and Shiller (1988) report a similar persistence measure of discount rate news, which they provide as evidence of long horizon return predictability.

Equation (1.13) provides the decomposition of fixed cost news into proportional cost news and dividend news. Using the decomposition, the fixed cost news variance may be written:

$$\text{var}(\eta_{c,t}) = \text{var}(\eta_{d,t}) + \text{var}(\bar{\eta}_{K,t}) + 2(\eta_{d,t}, \bar{\eta}_{K,t}) \quad (1.20)$$

Is new information about future fixed costs primarily due to changes in expected future proportional costs or dividends? In order to determine the relative volatility contributions of the two fixed cost news components, I calculate the following three measures for the above decomposition. The first measure is the relative volatilities of the two components: $R_d^* \equiv \sigma(\eta_{d,t})/\sigma(\eta_{c,t})$ and $\hat{R}_K^* \equiv \sigma(\bar{\eta}_{K,t})/\sigma(\eta_{c,t})$. The second measure is the respective regression coefficients obtained by regressing dividend and scaled proportional cost news on fixed cost news: $B_d^* \equiv \text{cov}(\eta_{c,t}, \eta_{d,t})/\text{var}(\eta_{c,t})$ for dividend news and $\bar{B}_K^* \equiv \text{cov}(\eta_{c,t}, \bar{\eta}_{K,t})/\text{var}(\eta_{c,t})$ for proportional cost news.

### 1.3 Vector Autoregressions

Campbell and Shiller (1988) show that a vector autoregression (VAR) is a convenient way to implement the return and return variance decomposition and estimate the news series.
I assume that the data are generated by a first-order VAR model,

\[ z_t = a + \Gamma z_{t-1} + w_t, \]  

(1.21)

where \( z_t \) is a \( m \)-by-1 portfolio-specific vector of state variables describing a portfolio at time \( t \), \( a \) and \( \Gamma \) are, respectively, an \( m \)-by-1 vector and an \( m \)-by-\( m \) matrix of parameters, and \( w_t \) is an \( m \)-by-1 vector of i.i.d. shocks. This formulation is not restrictive and allows for higher-order representation by including additional lags in the vector of state variables. The VAR coefficient matrix may differ across portfolios but is assumed to be time invariant. The error vector \( w_t \) has portfolio dependent covariance matrix \( \Sigma \). Writing the VAR in companion form simplifies forecasting. For example, the change in expectations at time \( t \) of the state variables \( i \) periods ahead is \( E_t[z_{t+i}|z_t] - E_{t-1}[z_{t+i}|z_{t-1}] = \Gamma^i w_t \).

The state vector includes the log gross return and proportional cost, which are required for calculating the news components, and additional variables that aid in forecasting. The additional state variables included for their forecasting abilities are described in the next section.

I define \( e_j \) to be the \( j^{th} \) row of an appropriately sized identity matrix. The vector \( e_j \) extracts \( j^{th} \) variable of the vector \( z_t \). For instance, the unexpected net return at date \( t \) is \( \tilde{\nu}_t = (e_1 - e_2)w_t \), the unexpected contemporaneous proportional cost is approximately \( \Delta E_t K_t = e_2 w_t \), and the net discount rate news is calculated as follows:

\[
\eta_{\tilde{r},t} = \Delta E_t \sum_{i=1}^{\infty} \rho^i \tilde{r}_{t+i} = (e_1 - e_2) \sum_{i=1}^{\infty} \rho^i \Gamma^i w_t \\
= (e_1 - e_2) \Gamma (I - \rho \Gamma)^{-1} w_t \\
= (e_1 - e_2) \Lambda w_t, 
\]

(1.22)

where \( \Lambda \) is defined by \( \Lambda \equiv \rho \Gamma (I - \rho \Gamma)^{-1} \), a nonlinear function of the VAR coefficients. The term \( (I - \rho \Gamma)^{-1} \), which is equal to \( (I + \Lambda) \), gives more persistent variables a higher weight.
The derivation of proportional cost news is similar with the following resulting expression:

\[ \eta_{K,t} = e_2(I + \Lambda)w_t \]  \hspace{1cm} (1.23)
\[ \bar{\eta}_{K,t} = \frac{1}{\omega_c}e_2(I + \Lambda)w_t. \]  \hspace{1cm} (1.24)

As is the case in the Campbell (1991) implementation, dividend news is estimated as a residual:

\[ \eta_{d,t} = e_1(I + \Lambda)w_t \]  \hspace{1cm} (1.25)
\[ \bar{\eta}_{d,t} = \omega_d e_1(I + \Lambda)w_t. \]  \hspace{1cm} (1.26)

Fixed cost news is estimated using equation (1.13):

\[ \eta_{c,t} = (e_1 + \frac{1}{\omega_c}e_2)(I + \Lambda)w_t \]  \hspace{1cm} (1.27)
\[ \bar{\eta}_{c,t} = (\omega_c e_1 + e_2)(I + \Lambda)w_t. \]  \hspace{1cm} (1.28)

The above expressions may be used, along with the coefficient matrix \( \Gamma \) and the innovation covariance matrix \( \Sigma \), to estimate the desired news covariance matrix and other statistics of interest. For instance, the three previously introduced measures of fixed cost news’s influence on portfolio volatility are calculated using the following equations:

\[ R_c = \sqrt{\frac{(\omega_c e_1 + e_2)(I + \Lambda)\Sigma(I + \Lambda)'(\omega_c e_1 + e_2)'}{(e_1 - e_2)\Sigma(e_1 - e_2)'}} \]  \hspace{1cm} (1.29)
\[ B_c = \frac{(\omega_c e_1 + e_2)(I + \Lambda)\Sigma(e_1 - e_2)'}{(e_1 - e_2)\Sigma(e_1 - e_2)'} \]  \hspace{1cm} (1.30)
\[ P_c = \sqrt{\frac{(\omega_c e_1 + e_2)(I + \Lambda)\Sigma(I + \Lambda)'(\omega_c e_1 + e_2)'}{e_2 \Sigma e_2'}}. \]  \hspace{1cm} (1.31)

The VAR coefficient matrix \( \Gamma \) is estimated using OLS and is reported with its OLS standard errors. Robust standard errors and small-sample bias corrected estimates are obtained for all relevant statistics, such as the news’ covariances, using the delete-an-observation jackknife
The jackknife method was selected over the traditional bootstrap because the random sampling of errors often led to ill-conditioned \((I - \rho \Gamma)\) matrices.\(^5\) Details on the jackknife bias-correction and the estimation of jackknife standard errors are included in the Appendix.

\[\text{1.4 State Variables, Data Set, and Illiquidity Proxy}\]

As shown in the previous section, two variables are needed to estimate the three news series: log gross returns and proportional costs. Because the news series are changes in forecasts of their respective variables, variables that aid in forecasting should also be included in the VAR. For this reason, the state vector also includes the log yield and its two lags, lagged proportional costs, and the log first difference of monthly turnover.

The price-dividend’s (or its inverse) forecasting ability is well-established in the return forecasting literature. The Campbell and Shiller (1988) log-linear approximation of the price-dividend ratio shows why the log yield would have predictive power for returns. Various forms of the dividend-price ratio are typically included in the return decomposition framework. Campbell (1991) includes the ratio of dividends paid over the previous year to the current stock price. Campbell and Ammer (1993) include the log of the Campbell (1991) measure. Campbell and Vuolteenaho (2004) choose to include the ratio of earnings paid over the previous year to the current stock price. Chen and Zhao (2006) investigate the robustness of the return decomposition to the choice of the price-earnings vs. price-dividend ratio. They estimate the model with the PE measure employed by Campbell and Vuolteenaho (2004) and a smoothed yield and find that the results of the two specifications differ sub-


\(^6\)Ill-conditioned matrices occurred approximately 20 percent of the time. Rather than set an arbitrary rule for removing these outliers, I choose the more structured jackknife approach, which eliminates the problem.
stantially, which they partially attribute to the fact that the unit root test is rejected for the yield, but not for the PE ratio. In order to avoid the persistence problems associated with smoothed variables, I do not smooth dividends like the above papers. Instead, I include the contemporaneous log yield and two additional lags of the log yield to account for the quarterly seasonality in dividend payout.

The proportional cost is a required state variable in my decompositions, but it also helps to forecast returns. Two papers that investigate the time-series return-liquidity relationship, Amihud (2002) and Jones (2002), report that high returns are predicted when liquidity is low. Acharya and Pedersen (2005) and Pastor and Stambaugh (2003) predict illiquidity using an AR(2) specification. In their spirit and because I implement the Acharya and Pedersen (2005) proportional cost proxy, an additional lag of proportional costs is included in the state vector.

The final state variable is the log first difference of monthly turnover, \( \Delta \psi_t = \psi_t - \psi_{t-1} \). Jones (2002) reports that high turnover predicts low returns and low trading costs. Because turnover is persistent, I choose to include its logged first difference instead to ensure stationarity.

The price, return, dividend, and volume data are from the Center for Research in Security Prices (CRSP) tape from July 1, 1962 until December 31, 2002 for all common shares (share codes 10 and 11) listed on the NYSE and AMEX (data with exchange codes 3 and 33 are omitted). The risk-free rate used to calculate excess returns are the one month Treasury-bill returns supplied by Ibbotson Associates. Monthly turnover is calculated as the monthly trading volume divided by the number of shares outstanding at the end of the current month. CRSP reports returns both inclusive and exclusive of dividend payout over the period. The log yield is calculated as the log of the equal-weight average single-period yield of the portfolio assets.
1.4.1 A Measure of Trading Costs

Unfortunately, illiquidity is not an observable variable and the data that measures certain aspects of liquidity, such as the bid-ask spread, are limited. For example, CRSP provides bid and ask prices for stocks listed on the NASDAQ only after 1982. As a result, researchers have proposed and investigated a number of liquidity proxies.\(^7\) I choose to proxy for trading costs using the methods of Acharya and Pedersen (2005), which are based on the following Amihud (2002) measure of illiquidity:

\[
\text{ILLIQ}_t^i = \frac{1}{\text{Days}_t^i} \sum_{d=1}^{\text{Days}_t^i} \frac{|\tilde{R}_{id}|}{V_{id}^i},
\]

where \(\text{Days}_t^i\) is the number of valid observation days for asset \(i\) in month \(t\) and \(V_{id}^i\) is the dollar volume in millions on day \(d\) in month \(t\). Given the specification of \(\text{ILLIQ}_t^i\), one may be concerned that \(\text{ILLIQ}_t^i\) is proxying for return volatility. Amihud (2002) shows that the correlation between his illiquidity measure and asset volatility is low, approximately 0.22. Also, Acharya and Pedersen (2005) test for this effect as a robustness check and find that including volatility as a state variable in their estimation does not significantly change their results. For a more complete discussion on the merits of using \(\text{ILLIQ}_t^i\) in this setting, see Acharya and Pedersen (2005).

The Amihud (2002) illiquidity measure cannot be used without modification because \(\text{ILLIQ}_t^i\) is measured in percent per dollar whereas my model requires a proportional measure of transactions costs in terms of dollar cost per dollar invested. This problem is addressed by relating \(\text{ILLIQ}_t^i\) to \(K_t^i\), the proportional monthly trading costs:

\[
K_t^i = 0.048 \min \left( 0.0025 + 0.0030\text{ILLIQ}_t^i\Lambda_{t-12}, 0.30 \right),
\]

where \(\Lambda_{t-12}\) is the ratio of the capitalizations of the market portfolio at the end of month.

\(^7\)See Goyenko, Holden, Lundblad, and Trzcinka (2005) for a comprehensive overview of some of the common proxies for liquidity.
$t-12$ and of the market portfolio at the end of July 1962. The original Acharya and Pedersen (2005) specification uses $\Lambda_{t-1}$ rather than $\Lambda_{t-12}$. A higher lag is implemented in this paper to avoid a spurious relationship in the VAR where the log dividend-price ratio is included in the state vector. Acharya and Pedersen (2005) select the coefficients 0.0025 and 0.0030 to match the cross-sectional distribution of $K_i^t$ for size-decile ranked portfolios to the effective half-spread reported by Chalmers and Kadlec (1998)\(^8\). Because I consider the cost incurred each period rather than that paid specifically at liquidation, I multiply by 0.048, the equal-weight average monthly turnover in my sample, to transform the full effective spread into a monthly amortized measure.\(^9\) Proportional trading costs are capped at 30 percent per trade to limit the impact of extreme observations of $ILLIQ_i^t$.

### 1.4.2 Inclusion Requirements

My inclusion requirements closely resemble those imposed by Amihud (2002), Pastor and Stambaugh (2003), and Acharya and Pedersen (2005). The following requirements are designed to eliminate stocks whose proportional cost proxies may behave poorly:

1. Stocks with end of previous year share prices of less than $5 or greater than $1000 are excluded.

2. Only stocks with more than 15 observations in a given month are included in the aggregate portfolio. An exception is made for the month of September 2001 because no stocks meet this requirement due to the six day market closure. For this month, to be included, a stock must have 15 observations.

3. The first and last partial month that a stock appears on the CRSP tape are excluded from the sample.

\(^8\)Chalmers and Kadlec (1998) report that the mean effective spread is 1.11%.

\(^9\)Acharya and Pedersen (2005) multiply by 0.034, which is the value-weighted monthly turnover and corresponds to a holding period of 29 months. Because I investigate the equal-weight market portfolio, I use the equal-weighted monthly turnover, which I estimate to be 0.048.
4. Stocks must have at least 120 observations in the previous year.

The initial CRSP sample of ordinary stocks on the NYSE and AMEX between January 1964 and December 2001 includes 1,009,698 observations. 784,908 observations survive the first three requirements. The fourth requirement is designed to ensure a suitable amount of data is available for estimating the level and volatility of individual stock illiquidity for the cross-sectional portfolio analysis and eliminates an additional 27,056 observations. The resulting data set includes 757,852 observations.

1.4.3 Portfolio Formation

An equal-weight market portfolio is formed each month from January 1964 to December 2002. Acharya and Pedersen (2005) focus on equal-weighted averages for the market portfolio. They argue persuasively that equal weights compensate for the over-representation of liquid assets in the CRSP sample. Because my sample and investigation is similar to and my results are compared against theirs, I too will focus my empirical work on the equal-weighted market portfolio.

In addition to the market portfolio, for each year-end beginning 1963 and ending 2001, eligible stocks are sorted into 3 portfolios according to criteria listed below. Portfolio characteristics for the 12 post-ranking months are linked across years to form a single time series for each quintile. For the quintile-ranked portfolios, all state variables, $z_t$, are calculated as value-weighted averages. The breakpoints for the sort are calculated using all eligible stocks so that each portfolio has approximately the same number of stocks when formed. The sort criteria is listed below:

- **Size** – Stocks are sorted by their mean market capitalizations in December of the previous year.

- **Proportional Cost** – Stocks are sorted by an annualized version of the Amihud (2002) illiquidity measure for each eligible stock where equation (1.32) is estimated over the entire year.
• **Turnover** – Stocks are sorted by their respective mean daily turnover calculated over the entire year.

The purpose of the above sorts is to investigate how liquidity news properties vary over the cross-section. For example, Stoll and Whaley (1983) report that small-cap stocks tend to be less liquid than large-cap stocks. Do they also have more liquidity risk? Acharya and Pedersen (2005) report that less liquid stocks are associated with higher levels of systematic liquidity risk. Does the finding hold when considering liquidity news risk? The cross-sectional analysis helps to answer these questions.

### 1.4.4 Calculating the News Coefficients

In Section 2, I emphasized that the weight applied to liquidity news is important over the cross-section. Recall that the liquidity news term is multiplied by the following coefficient:

$$
\omega_c = \frac{1}{D/C - 1}.
$$

(1.34)

The coefficient is sensitive to changes in the long-run dividend-cost ratio. To see this, consider the dividend cost ratios of the high and low portfolios ranked by market capitalization (Table 1.5). The average yield and proportional cost of the high portfolio is respectively 3.5776 and 0.1343 percent. The resulting coefficient is 0.04. The average yield and proportional cost of the low portfolio is respectively 2.0419 and 1.2332 percent. The resulting coefficient is 1.52, 39 times larger than that of the high portfolio.

Using these weights more than likely exaggerates the cross-sectional differences. At the same time, applying the same market weight to all portfolios is also suspect because the important cross-sectional differences are completely ignored. My solution is to average each portfolio’s yield and proportional cost with the market value prior to calculating the weight. This is a relatively conservative compromise that assumes the cost-dividend ratio of individual portfolios revert to that of the market over long horizons. Applying the compromise to the above example, I obtain coefficients of 0.11 and 0.55 for the portfolios. The high portfolio’s coefficient is now approximately 5 times smaller than that of the low portfolio.
Because of the sensitivity of the weight to the dividend-cost ratio and the noisiness of the Acharya and Pedersen (2005) proportional cost proxy implemented in this paper, I choose to use an alternative measure of long-term trading costs for calculating the coefficients. Chalmers and Kadlec (1998) report the amortized spread for effective spread decile-ranked portfolios. At the end of each year, I rank stocks by their annual Amihud (2002) illiquidity measure and assign to each stock the amortized spread reported by Chalmers and Kadlec (1998) for its respective decile-ranked portfolio. Each month, I calculate the equal-weighted average proportional cost for each portfolio. Finally, I use the time-series average proportional cost as the long run mean cost-price ratio.

1.5 Results of the VAR Decompositions

Table 1.2 reports the coefficient estimates for the VAR model for the equal-weight market portfolio formed over the period January 1964 to December 2001. Each estimate includes OLS standard errors in brackets and robust jackknife standard errors in parentheses. The first column is the intercept and the next seven columns are the coefficient estimates that form the coefficient matrix $\mathbf{\Gamma}$. The final two columns report the $R^2$ and $F$-statistic for each regression. The second panel reports the correlation matrix of the regression residuals.

The coefficient estimates are consistent with the return predictability literature. Consistent with the findings of Campbell and Shiller (1988), Campbell (1991), Campbell and Ammer (1993), Vuolteenaho (2002), Jones (2002), and Campbell and Vuolteenaho (2004), high returns and yields forecast high returns. High proportional costs positively predict the market return, consistent with Jones (2002), Amihud (2002), and Bekaert, Harvey, and Lundblad (2006). An increase in turnover forecasts higher returns, a result that is inconsistent with the findings of Jones (2002).\(^{10}\) The $R^2$ for the regression is 4.8 percent, which is consistent amongst the return predictability literature.

\(^{10}\)These differences can be attributed to the differences in our samples. Jones (2002) considers annual excess returns and his own constructed measurement of the bid-ask spread for the entire century.
Proportional costs are persistent and predictable. The $R^2$ of the regression is 69.5 percent. Like Jones (2002), I find that high returns predict low proportional costs. An increase in turnover forecasts high costs, a result which is consistent with the findings of Jones (2002).

Two interesting differences emerge between the results reported here and those obtained by Acharya and Pedersen (2005): (i) their AR(2) specification obtains an $R^2$ of 78 percent, which is higher than that obtained in this paper under the richer specification and (ii) the AR(2) coefficient in their specification is statistically significant whereas the AR(2) coefficient in my specification is not. The differences may be attributed to slight differences in the data filter, sample period, and proportional cost proxy specification.

Tables 1.3 and 1.4 report the variance decompositions implied by the VAR coefficients reported in Table 1.2. In order to implement the decomposition, values for $\omega_d$, $\omega_c$, and $\rho$ are required. Using the time-series mean of the portfolio’s monthly yield and Chalmers and Kadlec (1998) costs, the calibrated values are: $\omega_d = 1.242$ and $\omega_c = 0.242$. Consistent with the return decomposition literature, beginning with Campbell and Shiller (1988) and continuing with Campbell (1991), I set $\rho = 0.95^{1/12}$. Panel A reports the covariance and correlation matrix for the relevant news terms. Panel B relates the state vector shocks to the news terms through their correlations and the function that maps the state vector shocks to the individual news components. The functional mappings are normalized by multiplying by the respective state variable shock volatility and dividing by the unexpected contemporaneous net return volatility. The resulting estimate is the impact in standard deviations to contemporaneous returns of a single deviation shock to a particular state variable through the specified news series. Panel C provides the seven measures of liquidity’s contribution to portfolio risk. Robust jackknife standard errors are in parenthesis and small-sample bias corrected estimates using the jackknife procedure are in brackets.

### 1.5.1 Proportional Cost Decomposition

Table 1.3 reports the estimates for the previously defined statistics for the proportional cost decomposition. Discount rate news has the largest volatility of the three components, a result that is consistent with the Campbell (1991) decomposition. From $R_{\tilde{r}}$, we see
that discount rate news has approximately 107 percent of the variability of unexpected contemporaneous returns. According to $R_d$, dividend news is also an important component with 104 percent the variability of unexpected returns. As expected, liquidity’s contribution to portfolio volatility is the smallest of the three components; it’s volatility is approximately 1 percent of that of contemporaneous returns. The beta coefficients ($B$) take into account the covariation between the three news terms and paint a similar picture. Discount rate news is the largest contributor to portfolio variability with $B_r = -0.361$. Dividend news has about $7/4$ the impact on returns as discount rate news with $B_d = 0.638$. The covariation of proportional cost news with dividend and discount rate news virtually eliminates its impact on contemporaneous returns with $B_c = -0.001$.

Although proportional cost news is a small component of unexpected returns, its volatility is significantly larger than that of contemporaneous proportional costs. The persistence measure $P_K$ indicates that the revision in expected present and future proportional costs has approximately 5.2 times the volatility of contemporaneous proportional costs. From Panel B, we see that the correlation between illiquidity shocks and liquidity news is 0.784. If the VAR for proportional costs were limited to an AR(2) process, then by definition, the correlation between the contemporaneous shock and the news term would be unity. Apparently, the additional information provided by the other state variables is important for forecasting future proportional costs.

### 1.5.2 Fixed Cost Decomposition

Table 1.4 reports the statistics for the fixed cost decomposition. By definition, the role of discount rate news is identical in the two decompositions. Fixed cost news has a more significant impact on contemporaneous returns than proportional cost news. The $\hat{R}_c$ statistic shows that fixed cost news has approximately 24.6 percent the variability of contemporaneous returns versus the 1.2 percent for proportional cost news in the previous decomposition. The relationship $\eta_{c,t} = \eta_{d,t} + \hat{\eta}_K,t$ suggests that fixed cost news is primarily driven by new information about dividends. Table 1.3 shows that $\eta_{d,t}$ is substantially more volatile than $\eta_{K,t}$, even after dividing $\eta_{K,t}$ by $\omega_c = 0.242$. New information about dividends influences
prices and when proportional costs are relatively stable, the price change due to dividend news is the primary source of fixed cost news variability. As a measure of liquidity risk, fixed cost news is substantially more volatile than contemporaneous proportional costs: 105 times more volatile according to $\tilde{P}_c$. The fixed cost decomposition indicates that new information about future liquidity costs is an important source of risk for contemporaneous returns and that liquidity risk, as measured by fixed cost news risk, is significantly more important than the risk associated with unpredicted contemporaneous proportional costs.

### 1.5.3 Cross-Sectional Variation in the Variance Decompositions

A number of studies have investigated return predictability using the VAR decomposition initially proposed by Campbell and Shiller (1988). One study by Vuolteenaho (2002) estimates the variance decomposition over the cross-section to investigate how the properties of cash flow news and discount rate news differ across firms. Vuolteenaho (2002) forms ten size-ranked portfolios on an annual basis. Vuolteenaho’s approach assumes that all portfolios share the same coefficient matrix and that heterogeneity across portfolios is due to portfolio-specific innovation covariance matrices. This assumption eliminates the complication arising from invalidation of the infinite sum formulas. He notes that a result of the assumption is the possibility that heterogeneity in the decomposition may be an artifact of the imposed constraint. Imposing the same restriction is problematic in my sample. For example, some portfolios would have variances conditional on the state vector higher than unconditional variances. I choose to take a different approach and assume at the time of portfolio formation that a stock will always have the same transition matrix going forward. Therefore, I estimate portfolio dependent transition and innovation covariance matrices.

Each set of results for the various sorts are presented with sample statistics, the two variance decompositions, and the decomposition of fixed cost news into dividend and proportional cost news. The portfolio sample statistics include the expected gross returns, variance of gross returns, variance of the unexpected component of net returns, average market capitalization, yield, Chalmers and Kadlec (1998)’s amortized spread, the Acharya and Pedersen (2005) proportional cost measure, and turnover. For the variance decompositions, I include
the 14 statistics described earlier. Market capitalization is measured in billions of dollars and the remaining terms are annualized and reported as percentages. The rightmost column reports the estimates for the equal-weight market portfolio for ease of comparison.

Sort by Market Capitalization

Table 1.5 presents the results for the portfolios formed after sorting assets by their previous year-end market capitalizations. The sample statistics indicate that sorting by firm size leads to the usual results. Expected returns, return variance, and trading costs decrease with firm size, and the yield increases with firm size. Annual turnover only varies slightly across portfolios.

The same relationship between liquidity risk and firm size occurs under the two decompositions. The relative volatility of proportional cost news decreases with firm size, from 0.0560 for the smallest quintile to 0.0023 for the largest quintile. The relative volatility of fixed cost news also decreases with firm size, from 0.598 for the smallest quintile to 0.124 for the largest. For all five portfolios, fixed cost news appears to be primarily driven by dividend news. Panel D reports that the volatility of fixed cost news is slightly less than its dividend news component for all five portfolios. Proportional cost news’s contribution to fixed cost news variability decreases with firm size. For instance, proportional cost news has approximately 9.3 percent the volatility of fixed cost news for the smallest quintile and 1.8 percent the volatility of fixed cost news for the largest quintile.

Sort by Amihud (2002) Illiquidity Level

Table 1.6 reports the results for firms sorted by their illiquidity level. Due to the close inverse relationship between firm size and liquidity level, the cross-sectional variation of liquidity sorts almost perfectly mimics that of size sorts, only in the opposite direction. Approximately 67 percent of firms are placed in the same quintile-ranked portfolio when sorting on size or illiquidity level.

Unfortunately, the close relationship between firm size and firm liquidity makes it difficult to
determine which firm characteristic contributes to the cross-sectional variation of liquidity news. Panel A shows that firms with high costs have low market capitalization. The variance decomposition and measures of contribution also closely follow the results of the size sort. Like Acharya and Pedersen (2005), who show that illiquid securities tend to have high contemporaneous systematic liquidity risk, I find that illiquid securities also have high liquidity news risk as measured by the two decompositions.

Sort by Turnover

Table 1.7 presents the results for turnover sorted portfolios. The Amihud and Mendelson (1986) clientele effect predicts that high turnover stocks are more liquid. The reported annualized proportional costs for the five portfolios are consistent with the prediction. The liquid, high turnover portfolios have lower returns and yields. Although turnover and liquidity are linked and liquidity and market capitalization are related, there does not appear to be a relationship between turnover and size.

The turnover sort leads to an interesting set of results. The relative proportional cost volatility decreases in turnover, but the relative fixed cost volatility increases in turnover. The volatility of fixed and proportional cost news is, respectively, high and low for stocks that turn over most frequently. As Panel D reports, for these stocks, fixed cost news is primarily driven by dividend news. The volatility of fixed and proportional cost news is, respectively, low and high for stocks that turn over least frequently and proportional cost news has approximately half the volatility of fixed cost news. For these stocks, new information about proportional costs is a significant source of information about fixed costs.

1.5.4 Regime Shifts and Parameter Stability

In addition to the time-series fluctuations in expected and unexpected liquidity, occasionally the market undergoes large persistent shocks to the cost of trade. A recent example is the shift from fractional to decimal prices on January 29, 2001 for the NYSE and April 9, 2001 for the NASDAQ. Bessembinder (2002) reports that equal-weight average spread decreased
from 16.5 to 10.9 cents per share across all NYSE stocks and from 17.5 to 13.4 cents per share across all NASDAQ stocks post decimalization. Unfortunately, given the sample periods considered in this paper, testing for regime shifts associated with decimalization is infeasible.

Jones (2002) documents an earlier persistent shock to trading costs due to the commission deregulation that began in 1971 and continued until completion on May 1, 1975. Jones (2002) reports that beginning in 1976 commissions on the NYSE fell dramatically with a half-life of approximately seven to eight years. In this section, I test whether commission deregulation was associated with a shift in the contribution of liquidity news to portfolio volatility. An alternate approach would be to test for parameter stability with an unknown breakpoint using a method such as that suggested by Sowell (1996). The approach is more appropriate when either the time-series lacks a distinct and concrete event that may be associated with a shift in parameter values or the econometrician is unaware of the timing or existence of such an event. In this case, commission deregulation had a significant effect on trading costs and the specific event forms the basis of my test.

I begin by testing for VAR parameter stability across regimes. First, I consider the traditional Chow test for each dependant variable separately. With eight parameters and 453 observations, the critical value at the five and one percent levels for the F-distributed test statistics are, respectively, 1.96 and 2.55. The test statistics for the four regressors (return, proportional cost, log yield, and first difference of log turnover) are, respectively, 2.17, 1.06, 1.30 and 1.01. The test of parameter instability is only significant at the five percent level for the gross return. The second test simultaneously considers the system of four equations. Through two-stage GMM, I estimate the Likelihood Ratio test statistic for the hypothesis that the parameters pre-1976 are equal to those post-1976. The resulting test statistic of 30.86 is \( \chi^2 \) distributed with 32 degrees of freedom and has a p-value of 0.52. Hence, the joint test of parameter stability is not rejected at all standard confidence levels.

Because the VAR parameters appear to be stable over the considered breakpoint, I compare the volatilities of proportional cost news terms estimated over the respective period using the common VAR parameter estimates reported in Table 1.2. The annualized volatility of
proportional cost news over the entire sample period is 28.7 basis points. Prior to 1976 the annualized volatility is 26.5 basis points and subsequent to 1976 the volatility is 29.6 basis points. Due to the substantial accumulation of estimation error in the infinite sum formula that defines proportional cost news, the standard error of the individual estimates are considerable: 32 basis points for the first period and 24 basis points for the second. The standard error for the difference in volatilities is 40 basis points, which is associated with a confidence region 79 basis points wide. The confidence band is almost three times as wide as the larger estimate, making any meaningful comparison from a statistical perspective infeasible. Economically, however, the estimates between the two periods appear to be quite close and the influence of information about liquidity on return volatilities appears to be stable over time.

1.6 Pricing Liquidity Risk

The Acharya and Pedersen (2005) derived Liquidity-Adjusted Capital Asset Pricing Model (LACAPM) is a significant contribution that allows liquidity risk to be priced and asset returns to be explained as a function of their systematic liquidity risk. Their result is the solution to an overlapping generations equilibrium model with risk-averse agents and assets with stochastic liquidity levels. The unconditional expected net return of asset $i$ is

$$E(\tilde{r}_t^i) = E(r_t^i) + \lambda \frac{\text{cov}(\tilde{r}_t^i - E_{t-1}(\tilde{r}_t^i), \tilde{r}_t^m - E_{t-1}(\tilde{r}_t^m))}{\text{var}(\tilde{r}_t^m - E_{t-1}(\tilde{r}_t^m))}. \tag{1.35}$$

The LACAPM is a natural extension to the CAPM, exchanging the net return that agents care about in a world with liquidity costs for the gross return in the original CAPM. Acharya and Pedersen (2005) substitute in the standard decomposition of net return: $\tilde{r}_t^i \approx r_t^i - K_t^i$, which I refer to as the \textit{contemporaneous decomposition}, and obtain the following model for expected excess gross returns:

$$E(r_t^i - r_t^f) = E(K_t^i) + \lambda \beta_{r,r}^i + \lambda \beta_{c,c}^i - \lambda \beta_{c,r}^i - \lambda \beta_{r,c}^i \tag{1.36}$$
where
\[
\beta^i_{r,r} = \frac{\text{cov} \left( r^i_t - E_{t-1} \left( r^m_t \right), r^m_t - E_{t-1} \left( r^m_t \right) \right)}{\text{var} \left( r^m_t - E_{t-1} \left( r^m_t \right) - [K^m_t - E_{t-1} \left( K^m_t \right)] \right)}
\]  
(1.37)
is the systematic risk associated with covariation between an individual asset’s gross return with the market gross return,
\[
\beta^i_{c,c} = \frac{\text{cov} \left( K^i_t - E_{t-1} \left( K^m_t \right), K^m_t - E_{t-1} \left( K^m_t \right) \right)}{\text{var} \left( K^m_t - E_{t-1} \left( K^m_t \right) - [K^m_t - E_{t-1} \left( K^m_t \right)] \right)}
\]  
(1.38)
is the systematic risk due to comovement between an individual asset’s contemporaneous proportional liquidity level with that of the aggregate portfolio,
\[
\beta^i_{c,r} = \frac{\text{cov} \left( K^i_t - E_{t-1} \left( K^m_t \right), r^m_t - E_{t-1} \left( r^m_t \right) \right)}{\text{var} \left( r^m_t - E_{t-1} \left( r^m_t \right) - [K^m_t - E_{t-1} \left( K^m_t \right)] \right)}
\]  
(1.39)
represents the non-diversifiable risk due to an individual asset’s proportional liquidity sensitivity to the market return, and
\[
\beta^i_{r,c} = \frac{\text{cov} \left( r^i_t - E_{t-1} \left( r^m_t \right), K^m_t - E_{t-1} \left( K^m_t \right) \right)}{\text{var} \left( r^m_t - E_{t-1} \left( r^m_t \right) - [K^m_t - E_{t-1} \left( K^m_t \right)] \right)}
\]  
(1.40)
is the systematic risk that results from an individual asset’s return covarying with the aggregate proportional liquidity level. In their model, the expected excess net return \( \lambda = E(r^m_t - K^m_t - r^f_t) \) is the market risk premium.

This paper emphasizes through the return decomposition approach that the gross return contains a liquidity component and that the net return decomposition may be used to disentangle the liquidity component from gross returns. Note that equation (1.35) is equivalent to
\[
E(\tilde{r}^i_t) = E(r^f_t) + \lambda \frac{\text{cov}(\tilde{\nu}^i_t, \tilde{\nu}^m_t)}{\text{var}(\tilde{\nu}^m_t)}
\]  
(1.41)
\[
= E(r^f_t) + \lambda \tilde{\beta}.
\]  
(1.42)
The two-term decompositions, \( \tilde{\nu}^i_t \approx \eta^*_{K,t} - \eta^*_{K,t} \) for proportional costs and \( \tilde{\nu}^i_t \approx \eta^*_{c,t} - \eta^*_{c,t} \) for fixed costs, may be substituted into equation (1.42) to obtain two alternatively specified
LACAPMs. For the proportional cost decomposition, the four betas are defined as follows:

\[
\begin{align*}
\beta_{i,r,r}^i &= \frac{\text{cov}(\eta_{K,t}^i, \eta_{K,t}^m)}{\text{var}(\tilde{\nu}_t^m)} \\
\beta_{c,c}^i &= \frac{\text{cov}(\eta_{K,t}^i, \eta_{K,t}^m)}{\text{var}(\tilde{\nu}_t^m)} \\
\beta_{r,c}^i &= \frac{\text{cov}(\eta_{K,t}^i, \eta_{K,t}^m)}{\text{var}(\tilde{\nu}_t^m)} \\
\beta_{c,r}^i &= \frac{\text{cov}(\eta_{K,t}^i, \eta_{K,t}^m)}{\text{var}(\tilde{\nu}_t^m)}.
\end{align*}
\]

For the fixed cost decomposition, the four betas are defined to be

\[
\begin{align*}
\beta_{r,r}^i &= \frac{\text{cov}(\eta_{c,t}^i, \eta_{c,t}^m)}{\text{var}(\tilde{\nu}_t^m)} \\
\beta_{c,c}^i &= \frac{\text{cov}(\eta_{c,t}^i, \eta_{c,t}^m)}{\text{var}(\tilde{\nu}_t^m)} \\
\beta_{r,c}^i &= \frac{\text{cov}(\eta_{c,t}^i, \eta_{c,t}^m)}{\text{var}(\tilde{\nu}_t^m)} \\
\beta_{c,r}^i &= \frac{\text{cov}(\eta_{c,t}^i, \eta_{c,t}^m)}{\text{var}(\tilde{\nu}_t^m)}.
\end{align*}
\]

1.6.1 Estimating the Betas

In the contemporaneous decomposition, the four betas are estimated over the asset’s holding period, which Acharya and Pedersen (2005) estimate to be 29 months on average, and not over the sampling period as viewed by the econometrician. In their model, the transaction cost is paid every 29 months at asset liquidation and dividends are aggregated over the entire holding period. In order to estimate the betas with the 1 month sampling horizon, Acharya and Pedersen (2005) assume that returns are independent over time and proportional costs follow a martingale. In the Appendix, I show that small deviations from the two assumptions can lead to large differences in estimated betas due to error aggregation over the 29 month period. When returns follow an AR(1) process with coefficient \(\rho_i\) for portfolio \(i\) and \(\rho_m\) for the market portfolio and proportional costs follow an AR(1) process with coefficient \(\rho_i\) for portfolio \(i\) and \(\rho_m\) for the market portfolio, the four betas are calculated.
using the following equations for the relevant covariance terms:

\[
\begin{align*}
\text{cov}_t(K^i_{t+\tau}, K^m_{t+\tau}) &= \frac{1 - \rho^\tau_i}{1 - \rho^\tau_m} \text{cov}_t(K^i_{t+1}, K^m_{t+1}) \\
\text{cov}_t(r^i_{t,t+\tau}, r^m_{t,t+\tau}) &= \frac{\text{cov}_t(r^i_{t+1}, r^m_{t+1})}{(1 - \rho_i)(1 - \rho_m)} \left( \tau + \rho_i \rho_m \frac{1 - \rho^\tau_i}{1 - \rho_m} - \rho_i \frac{1 - \rho^\tau_m}{1 - \rho_m} - \rho_m \frac{1 - \rho^\tau_i}{1 - \rho_i} \right) \\
\text{cov}_t(K^i_{t+\tau}, r^m_{t,t+\tau}) &= \left( \frac{1 - \rho^\tau_i}{1 - \rho_m} - \frac{\rho_m}{1 - \rho_m} \frac{1 - \rho^\tau_i}{1 - \rho_m} \right) \text{cov}_t(K^i_{t+1}, r^m_{t+1}) \\
\text{cov}_t(r^i_{t,t+\tau}, K^M_{t+\tau}) &= \left( \frac{1 - \rho^\tau_i}{1 - \rho_i} - \frac{\rho_i}{1 - \rho_i} \frac{1 - \rho^\tau_i}{1 - \rho_i} \right) \text{cov}_t(r^i_{t+1}, K^m_{t+1}) \\
\text{var}_t(r^m_{t,t+\tau} - K^m_{t+\tau}) &= \frac{\text{var}_t(r^m_{t+1})}{(1 - \rho_m)^2} \left( \tau + \rho_m \frac{1 - \rho^2_m}{1 - \rho_m} - 2 \rho_m \frac{1 - \rho^\tau_m}{1 - \rho_m} + \frac{1 - \rho^2_m}{1 - \rho_m} \text{var}_t(K^m_{t+\tau}) \\
&\quad - 2 \left( \frac{1 - \rho^\tau_m}{1 - \rho_m} - \frac{\rho_m}{1 - \rho_m} \frac{1 - \rho^\tau_m}{1 - \rho_m} \right) \text{cov}_t(r^m_{t+1}, K^m_{t+1}) \right) \\
\end{align*}
\]  

where \( r_{t,t+\tau} \) is the cumulative log return over the \( \tau \) periods from time \( t \) to \( t + \tau \). For instance, when returns follow an AR(1) process with coefficient 0.167 and proportional costs follow an AR(1) process with coefficient 0.886, \( \beta^i_{c,c} \) is estimated to be 6.3 times larger than its true value and \( \beta^i_{c,r} \) and \( \beta^i_{r,c} \) are estimated to be 3.08 times larger than their true values.

As detailed in Section 2, I take a different approach for the proportional and fixed cost decompositions. Rather than aggregate dividends to be distributed at the end of the average holding period, I assume that at each sampling period, investors bear a cost for holding illiquid assets. Because, by assumption, costs are incurred and dividends are distributed at the sampling frequency, I am able to estimate the covariances and variances in equations (1.43) and (1.44) at the monthly frequency.

Tables 1.8 through 1.10 present the descriptive statistics for the odd-numbered value-weighted portfolios after sorting assets into 25 portfolios by their annual Amihud (2002) measures for the contemporaneous, proportional, and fixed cost decompositions respectively. For the contemporaneous decomposition, the four betas are estimated using equations (1.37) through (1.40) and (1.45) through (1.49). For the proportional and fixed cost decompositions, the betas are computed using equations (1.43) and (1.44) and the monthly estimated
news series for the 25 portfolios and the aggregate portfolio. In addition to the four betas, 
the aggregate net return beta ($\beta^{\text{net}} = \beta_{r,r}^i + \beta_{c,c}^i - \beta_{r,c}^i - \beta_{c,r}^i$) is reported for each port-
folio. All statistics are estimated over the period 1964 to 1999 so that my sample period 
matches that of Acharya and Pedersen (2005). The reported estimates are small-sample 
bias corrected and the standard errors are computed using the jackknife procedure and the 
pre-estimation of the coefficient matrix and residual covariance matrix are taken into ac-
count. For each portfolio, I also report the annualized time-series average proportional cost, 
net return, turnover, yield, and market capitalization as well as the time-series volatility of 
the portfolio’s proportional cost and net return.

Sorting assets on previous illiquidity levels results in portfolios with monotonically in-
creasing illiquidity, providing further evidence of liquidity’s persistence. The character-
istics of the portfolios are similar to those of the illiquidity sorted portfolios reported by 
Acharya and Pedersen (2005) in their Table 1, with the differences attributable to the slight 
difference in our inclusion requirements. Highly illiquid portfolios have high net returns, 
a relationship that is well documented in the literature. Yield and market capitalization 
monotonically increase with liquidity. Aside from portfolios 1 and 25, the inverse relation-
ship between illiquidity and turnover is consistent with the Amihud and Mendelson (1986) 
clientele effect.

Sorting on past illiquidity levels produces clear trends in the betas. The net return beta 
increases in illiquidity for all three decompositions. According to the betas reported in 
Table 1.8, adjusting for return’s momentum and liquidity’s mean reversion significantly 
alters the estimates. Comparing the liquidity betas reported in table 1.8 to those reported 
by Acharya and Pedersen (2005), we see that the estimates of $\beta_{c,c}$, $\beta_{c,r}$, and $\beta_{r,c}$ in this 
paper are, respectively, approximately one-tenth, one-third, and one-fourth the magnitude 

Tables 1.9 and 1.10 provide the same information for the betas associated with the propor-
tional and fixed cost decompositions. By construction, the net return beta for the fixed and 
proportional cost decompositions are equal. The magnitudes of the liquidity betas in the 
fixed cost decomposition are significantly larger than those obtained in the proportional cost
and contemporaneous decompositions, reflecting the larger variability in fixed cost news. All three decompositions indicate that the largest source of systematic liquidity risk is the covariation between an asset’s illiquidity and the market return. As Acharya and Pedersen (2005) point out, for illiquid stocks, a negative return coincides with an increase in liquidation costs. To make matters worse, from the proportional cost decomposition, we see that the price impact of changes in forecasted liquidity levels exacerbates the situation; $\beta_{c,r}$ is approximately three times larger in the proportional cost decomposition than in the contemporaneous decomposition. Not only does negative news about dividends or positive news about discount rates coincide with increases in illiquidity costs, but it also coincides with increases in expected illiquidity costs which further drives down the price.

1.6.2 The Liquidity Risk–Return Relationship

I closely follow the methods of Acharya and Pedersen (2005) to estimate the market price of liquidity risk. Again, I report the small-sample bias corrected estimates and standard errors calculated using the jackknife procedure. The reported standard errors take into account the pre-estimation of the betas, VAR coefficient matrix, and residual covariance matrix.

In addition to the 25 value-weighted portfolios formed by ranking stocks by their annual ILLIQ measures, Acharya and Pedersen (2005) also form 25 value-weighted portfolios by ranking stocks on a measure they denote $\sigma($illiquidity$)$, which is the volatility of daily ILLIQ values estimated over the previous year. The reason for forming portfolios in this manner is to consider portfolios that differ in their liquidity attributes. Tables 1.8 through 1.10 demonstrate that sorting by liquidity levels generates portfolios that differ in their liquidity attributes. I verify that sorting by $\sigma($illiquidity$)$ also produces portfolios that satisfy this requirement.

To estimate the risk premia, I consider the following specification with a variety of restrictions:

$$E(r_{it}^i - r_{ft}^f) = \alpha + \kappa E[K_t^i] + \lambda \beta_i^i + \lambda^* \beta^*i + \lambda_{r,r} \beta_{r,r}^i + \lambda_{r,c} \beta_{r,c}^i + \lambda_{c,r} \beta_{c,r}^i + \lambda_c \beta_c^i,$$

(1.50)
where \( \alpha \) allows for a nonzero intercept, \( \kappa \) is an estimate of the illiquidity premium as a multiple of proportional illiquidity costs, \( \beta \) is the CAPM beta, \( \beta^* \) is the net return liquidity-adjusted CAPM beta, and \( \beta_c^i \equiv \beta_{c,c}^i - \beta_{c,r}^i - \beta_{r,c}^i \) is a combined measure of the three components of systematic liquidity risk. The Acharya and Pedersen (2005) LACAPM is equation (1.50) with the following restrictions: \( \alpha = \lambda = \lambda_{r,r} = \lambda_{r,c} = \lambda_{c,r} = \lambda_{c,c} = \lambda_c = 0 \) and \( \kappa = 1 \).

Tables 1.11 through 1.13 report the estimated coefficients for both sets of sorted portfolios for the contemporaneous, proportional, and fixed cost decompositions respectively. The results for the illiquidity ranked portfolios are reported in Panel A and for the \( \sigma(\text{illiquidity}) \) ranked portfolios in Panel B.

Line 1 estimates the standard CAPM with betas calculated conditional on the state vector. The second line estimates the LACAPM with restricted liquidity premium and the third line removes the restriction on the liquidity premium. Lines 4 and 5 estimate the LACAPM allowing the market price of liquidity risk to differ from the market price of non-liquidity risk, but restricts the three sources of liquidity risk to have the same price of risk. The fourth line includes a restriction on the liquidity premium, which the fifth line removes. Finally, lines 6 and 7 remove the restriction that requires the three liquidity betas to share the same market price of risk. Line 6 includes a restricted liquidity premium and line 7 is the completely unrestricted model.

I begin the discussion with the contemporaneous decomposition that adjusts the betas for return momentum and liquidity mean reversion. As Acharya and Pedersen (2005) report in their empirical analysis, the restricted form of the LACAPM estimated in the second line provides an improvement in explanatory power over the standard CAPM. Table 1.11 shows the adjusted \( R^2 \) increasing from 55.0 percent for CAPM to 68.4 percent for the restricted LACAPM. Acharya and Pedersen (2005) report the adjusted \( R^2 \) to be 63.8 percent for CAPM and 73.2 percent for the restricted LACAPM. The overall improvement in fit is similar in the two different specifications. Removing the restriction on the liquidity premium, Acharya and Pedersen (2005) report an increase in the adjusted \( R^2 \) to 80.9 percent. Line 3 reports that the adjusted \( R^2 \) increases to 82.0 percent, a significant improvement.
from the restricted case. The specification estimated in line 4 allows the market price of contemporaneous systematic liquidity risk to differ from gross return risk. I report the market price of liquidity risk to be approximately 23 percent larger than that of gross return risk; the difference is not statistically significant. Acharya and Pedersen (2005) estimate the market price of liquidity risk to be approximately 277 percent larger than that of gross return risk, a statistically significant result. The primary finding of my analysis is that the explanatory power of the LACAPM is higher than that of CAPM and the market price of systematic contemporaneous liquidity risk is not statistically different from gross return risk. The fifth line releases the restricted liquidity premium, but the effects of over-fitting and multi-collinearity are apparent. The alpha is estimated to be almost 10 percent per year and the estimated market prices of risk are negative: -9 percent per year for gross return risk and -3114 percent per year for liquidity risk.

Table 1.12 reports the results for the proportional cost decomposition. There are two primary differences between the estimation of the market price of risk in the proportional cost decomposition and the contemporaneous decomposition considered above. First, the contemporaneous decomposition implicitly restricts the market price of the liquidity risk associated with new information about future expected liquidity to be the same as that of dividend and discount rate news. Second, in the contemporaneous decomposition of Acharya and Pedersen (2005), investors are only concerned with the liquidity level 29 months in the future when they liquidate their portfolio. In my specification, investors monetize the costs of holding an illiquid portfolio each period; they care if illiquidity is higher next month, even if they will not sell the asset for another 28 months. From lines 1 through 3, we see that the LACAPM provides an improvement in fit over the CAPM. The CAPM obtains an adjusted $R^2$ of 55.0 percent. When the liquidity premium is restricted, the adjusted $R^2$ of the LACAPM is 60.0 percent. Relaxing the restriction on the liquidity premium, the explanatory power increases to an adjusted $R^2$ of 79.4 percent. When the liquidity premium is restricted and the liquidity risk premium is unrestricted, the improvement in fit over the restricted liquidity risk premium model is slight with an adjusted $R^2$ of 60.9 percent. I report the market price of liquidity risk to be approximately 230 percent
higher than non-liquidity risk, a statistically significant result. As is the case in the contemporaneous decomposition, the unrestricted liquidity premium and unrestricted liquidity risk premium specification has a higher adjusted $R^2$ but multi-collinearity appears to be problematic. Setting aside the specifications with apparent over-fitting, the restricted liquidity risk premium model provides the best fit, suggesting that the market price of liquidity risk is not statistically different than that of non-liquidity risk.

Table 1.13 replicates the analysis for the fixed cost decomposition. By definition, the first three regressions are identical for the fixed and proportional cost decompositions. As is the case for the previous specifications, the fourth line estimates the LACAPM with a restricted liquidity premium and an unrestricted liquidity risk premium. I report an improvement in explanatory power over the restricted liquidity risk premium model, with an increase in adjusted $R^2$ from 60.0 to 64.9 percent. The market price of liquidity risk is estimated to be a statistically significant 77 percent higher than non-liquidity risk. Caution is needed, however, when interpreting the result. Fixed cost news is almost perfectly correlated with dividend news because fluctuations in the proportional cost are not enough to offset fluctuations in price. Hence, it is not clear whether the higher price of risk is attributed to fixed cost news or dividend news. Campbell and Vuolteenaho (2004) estimate a two-beta CAPM where dividend news and discount rate news have different prices of risk. They estimate the price of dividend risk to be higher than that of discount rate risk. The results in line 4 of Table 1.13 are consistent with their findings. Unfortunately, because of the close relationship between dividend news and fixed cost news, as evidenced by a correlation of 0.9993, identification is likely an insurmountable hurdle.

1.6.3 Economic Significance

So far, this section has focused on the explanatory power and the estimated market price of risk for the adjusted contemporaneous decomposition and the fixed and proportional cost decompositions within the LACAPM framework. The LACAPM may also be used to estimate how much of the expected return is attributable to each systematic risk component. In order to estimate the premia for the different liquidity components, I use the estimates
reported in line 3, where the market price of risk is restricted to be the same for the four betas and the liquidity risk premium is unrestricted. Line 3 is selected over line 4 because its adjusted $R^2$ is higher in each of the three specifications. Line 5 is rejected because of the apparent multi-collinearity problems that result in nonsensical estimates.

The model implied liquidity premium between portfolios 25 and 1 is $\kappa (E(K^{25}) - E(K^1))$. The difference in proportional costs between the two portfolios is 2.97 percent. For the contemporaneous decomposition, the estimated liquidity premium is 3.1 percent. The proportional and fixed cost decompositions have the same estimated liquidity premium of 4.1 percent annually. The estimates are consistent with the 3.5 percent annual liquidity premium reported by Acharya and Pedersen (2005).

I consider the estimated premium associated with the four sources of systematic risk according to the LACAPM. Using the estimated market price of risk in the contemporaneous decomposition of 71 basis points per month, the total risk premium is computed to be 4.1 percent per year. The systematic risk captured by $\beta_{r,r}$ generates an annual premium of 3.9 percent. The remaining 0.2 percent premium is for the comovement of an assets contemporaneous liquidity and the market return. That this source of liquidity risk is the most important of the three is consistent with the findings of Acharya and Pedersen (2005). The difference in magnitude – they estimate the premium to be 0.8 percent – is due to the adjustment for return momentum and liquidity mean reversion. The total liquidity related premium accounting for variation in the level of liquidity as well as systematic liquidity risk is 3.3 percent per year. Acharya and Pedersen (2005) estimate the overall effect to be 4.6 percent per year.

The proportional and fixed decompositions provide the same estimate for the total risk premium by construction, 3.3 percent per year. The difference between the two decompositions is in how the premium is divided among the different sources of risk. The proportional cost decomposition allocates 2.7 percent per year to non-liquidity risk (dividend and discount rate news risk) and 0.6 percent to the risk associated with the covariation between an asset’s liquidity news and the markets non-liquidity news. The fixed cost decomposition allocates 11.3 percent per year to non-liquidity risk and -8.1 percent per year to liquidity risk. The
magnitude is directly related to the higher volatility of fixed cost news versus proportional cost news. The negative sign is because fixed cost news is positively related to contemporaneous net returns through its tight link with dividend news. Although, according to the decomposition, new information about higher fixed costs has a negative influence on price, the positive liquidity news occurs when dividend news is positive.

Acharya and Pedersen (2005) write:

“The collinearity between liquidity and liquidity risk implies that the most robust number is their overall effect. Further, our results suggest that studies that focus on the separate effect of liquidity (or liquidity risk) can possibly be reinterpreted as providing an estimate of the overall effect of liquidity and liquidity risk.”

The quote appears to be relevant in the analysis presented here. The overall effect of liquidity in their analysis and the alternate contemporaneous decomposition investigated in this paper is similar. However, they attribute the majority of the premium to risk whereas I report the bulk of the premium is due to the level. The same phenomenon is found when comparing the contemporaneous decomposition to the proportional cost decomposition. The fixed cost decomposition, however, does not follow the logic because the collinearity is between liquidity risk and non-liquidity risk.

1.7 Conclusion

I derive two extensions of the Campbell (1991) unexpected return decompositions in order to include a liquidity component and investigate how new information about future liquidity levels influences portfolio volatility. The first extension includes an additional news term that represents changes in expected future proportional costs. The second extension’s additional news term represents changes in expected future fixed costs.

I estimate the news series for the two decompositions for the aggregate portfolio and over the cross-section. The primary findings are that proportional cost news and fixed cost
news have, respectively, 5.2 and 105 times the volatility of unexpected contemporaneous proportional costs. Fixed cost news has approximately 25 percent the volatility of contemporaneous returns and proportional cost news has approximately 1 percent the volatility of contemporaneous returns. Thus, fixed cost news appears to be an economically significant contributor to portfolio risk. Proportional cost news, on the other hand, is not a significant contributor to portfolio volatility. The impact of liquidity news risk varies over the cross section. Both measures of liquidity news risk have increasing volatilities as firm size decreases and illiquidity increases. Stocks that turn over more frequently have less volatile proportional cost news and more volatile fixed cost news.

Acharya and Pedersen (2005) propose and estimate a Liquidity-Adjusted CAPM that includes four betas that capture different elements of systematic risk associated with gross returns and contemporaneous proportional costs. They provide evidence that the LACAPM has higher explanatory power than the traditional CAPM and that the market price of systematic liquidity risk is statistically higher than that of systematic gross return risk. Their estimation procedure assumes that returns are independent over time and costs are a martingale. Relaxing the two assumptions and accounting for return momentum and liquidity mean reversion, I find that the market price of liquidity risk is no longer statistically different than gross return risk.

I investigate the explanatory power and pricing of systematic liquidity news risk under the two alternate specifications obtained through the net return decompositions. Like contemporaneous liquidity risk, systematic liquidity news risk helps to explain the cross-section of expected stock returns. The premium paid for systematic proportional cost news accounts for a 0.6 percent difference in annual expected returns between the least liquid and most liquid stocks. Comparing the result to the premium for systematic contemporaneous liquidity risk, we learn that the majority of the liquidity risk premium is due to new information about future liquidity and not unexpected contemporaneous liquidity.
Table 1.1: This table reports the sample statistics for the VAR state variables estimated for the period January 1963 to December 2001 for the equal-weight market portfolio. The included state variables are the log market gross return \(r_m\), proportional costs \(K\), log dividend yield \(y\), the first difference of log monthly turnover \(\Delta \psi\), one lag of proportional costs, and two lags of log dividend yield. The second panel reports the correlation and first-order autocorrelation of the series.

### Descriptive Statistics of the VAR State Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Median</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
<th>Autocorr.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(r_m)</td>
<td>1.079</td>
<td>1.468</td>
<td>5.417</td>
<td>−31.347</td>
<td>20.877</td>
<td>0.166</td>
</tr>
<tr>
<td>(K)</td>
<td>0.052</td>
<td>0.046</td>
<td>0.022</td>
<td>0.018</td>
<td>0.127</td>
<td>0.805</td>
</tr>
<tr>
<td>(y)</td>
<td>−3.669</td>
<td>−3.662</td>
<td>0.460</td>
<td>−4.754</td>
<td>−2.642</td>
<td>0.378</td>
</tr>
<tr>
<td>(\Delta \psi)</td>
<td>0.003</td>
<td>0.011</td>
<td>0.170</td>
<td>−0.600</td>
<td>0.658</td>
<td>−0.273</td>
</tr>
</tbody>
</table>

### Shock Correlation Matrix

<table>
<thead>
<tr>
<th>Correlations</th>
<th>(r_m)</th>
<th>(K_{t+1})</th>
<th>(y_{t+1})</th>
<th>(\Delta \psi_{t+1})</th>
<th>(K_t)</th>
<th>(y_t)</th>
<th>(y_{t-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(r_m)</td>
<td>1.000</td>
<td>−0.074</td>
<td>0.048</td>
<td>0.329</td>
<td>0.020</td>
<td>0.113</td>
<td>0.095</td>
</tr>
<tr>
<td>(K_{t+1})</td>
<td>−0.074</td>
<td>1.000</td>
<td>−0.213</td>
<td>0.044</td>
<td>0.805</td>
<td>−0.184</td>
<td>−0.110</td>
</tr>
<tr>
<td>(y_{t+1})</td>
<td>0.048</td>
<td>−0.213</td>
<td>1.000</td>
<td>−0.026</td>
<td>−0.087</td>
<td>0.378</td>
<td>0.383</td>
</tr>
<tr>
<td>(\Delta \psi_{t+1})</td>
<td>0.329</td>
<td>0.044</td>
<td>−0.026</td>
<td>1.000</td>
<td>−0.028</td>
<td>0.131</td>
<td>0.027</td>
</tr>
<tr>
<td>(K_t)</td>
<td>0.020</td>
<td>0.805</td>
<td>−0.087</td>
<td>−0.028</td>
<td>1.000</td>
<td>−0.213</td>
<td>−0.184</td>
</tr>
<tr>
<td>(y_t)</td>
<td>0.113</td>
<td>−0.184</td>
<td>0.378</td>
<td>0.131</td>
<td>−0.213</td>
<td>1.000</td>
<td>0.373</td>
</tr>
<tr>
<td>(y_{t-1})</td>
<td>0.095</td>
<td>−0.110</td>
<td>0.383</td>
<td>0.927</td>
<td>−0.184</td>
<td>0.373</td>
<td>1.000</td>
</tr>
<tr>
<td>(r_m)</td>
<td>0.166</td>
<td>−0.197</td>
<td>0.003</td>
<td>0.075</td>
<td>−0.074</td>
<td>0.051</td>
<td>0.113</td>
</tr>
<tr>
<td>(K_t)</td>
<td>0.020</td>
<td>0.805</td>
<td>−0.087</td>
<td>−0.028</td>
<td>1.000</td>
<td>−0.213</td>
<td>−0.184</td>
</tr>
<tr>
<td>(y_t)</td>
<td>0.113</td>
<td>−0.184</td>
<td>0.378</td>
<td>0.131</td>
<td>−0.213</td>
<td>1.000</td>
<td>0.373</td>
</tr>
<tr>
<td>(\Delta \psi_{t+1})</td>
<td>0.108</td>
<td>0.056</td>
<td>−0.059</td>
<td>−0.273</td>
<td>0.044</td>
<td>−0.031</td>
<td>0.136</td>
</tr>
<tr>
<td>(K_{t-1})</td>
<td>0.009</td>
<td>0.653</td>
<td>−0.147</td>
<td>0.064</td>
<td>0.806</td>
<td>−0.085</td>
<td>−0.213</td>
</tr>
<tr>
<td>(y_{t-1})</td>
<td>0.095</td>
<td>−0.110</td>
<td>0.383</td>
<td>0.027</td>
<td>−0.184</td>
<td>0.373</td>
<td>1.000</td>
</tr>
<tr>
<td>(y_{t-2})</td>
<td>0.009</td>
<td>−0.217</td>
<td>0.940</td>
<td>−0.062</td>
<td>−0.111</td>
<td>0.380</td>
<td>0.372</td>
</tr>
</tbody>
</table>
Table 1.2: This table reports the OLS parameter estimates for the equal-weight market portfolio for a monthly first-order VAR model including a constant, the log market gross return ($r^m$), proportional cost ($K$), log dividend yield ($y$), the first difference of log monthly turnover ($\Delta \psi$), one lag of proportional costs, and two lags of log dividend yield. The first eight columns report coefficients for the appropriate explanatory variable. The final two columns report the $R^2$ and $F$ statistics. OLS standard errors are in brackets below their respective parameter estimates and robust jackknife standard errors are in parentheses. In addition, the second panel reports the correlation matrix of the shocks with robust jackknife standard errors in parentheses.

**VAR Parameter Estimates: Jan 1964 – Dec 2001**

<table>
<thead>
<tr>
<th>Constant</th>
<th>$r^m_{t-1}$</th>
<th>100</th>
<th>$K_{t-1}$</th>
<th>$y_{t-1}$</th>
<th>$\Delta \psi_{t-1}$</th>
<th>100</th>
<th>$K_{t-2}$</th>
<th>$y_{t-2}$</th>
<th>$y_{t-3}$</th>
<th>$R^2$</th>
<th>$F$-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r^m_t$</td>
<td>0.066</td>
<td>0.149</td>
<td>0.277</td>
<td>0.013</td>
<td>0.017</td>
<td>-0.166</td>
<td>0.005</td>
<td>-0.000</td>
<td></td>
<td>4.8</td>
<td>3.21</td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.051)</td>
<td>(0.203)</td>
<td>(0.006)</td>
<td>(0.016)</td>
<td>(0.203)</td>
<td>(0.006)</td>
<td>(0.006)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100 - $K_t$</td>
<td>0.004</td>
<td>-0.065</td>
<td>0.767</td>
<td>0.001</td>
<td>0.009</td>
<td>0.035</td>
<td>0.005</td>
<td>-0.008</td>
<td></td>
<td>69.5</td>
<td>144.87</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.012)</td>
<td>(0.048)</td>
<td>(0.002)</td>
<td>(0.004)</td>
<td>(0.047)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$y_t$</td>
<td>-0.028</td>
<td>-0.666</td>
<td>-0.192</td>
<td>0.005</td>
<td>-0.171</td>
<td>0.893</td>
<td>0.061</td>
<td>0.935</td>
<td></td>
<td>89.8</td>
<td>561.60</td>
</tr>
<tr>
<td></td>
<td>(0.074)</td>
<td>(0.140)</td>
<td>(0.564)</td>
<td>(0.018)</td>
<td>(0.044)</td>
<td>(0.562)</td>
<td>(0.017)</td>
<td>(0.017)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta \psi_t$</td>
<td>0.067</td>
<td>0.550</td>
<td>-0.364</td>
<td>0.051</td>
<td>-0.323</td>
<td>0.685</td>
<td>0.023</td>
<td>-0.050</td>
<td></td>
<td>13.5</td>
<td>9.95</td>
</tr>
<tr>
<td></td>
<td>(0.079)</td>
<td>(0.151)</td>
<td>(0.607)</td>
<td>(0.019)</td>
<td>(0.048)</td>
<td>(0.605)</td>
<td>(0.019)</td>
<td>(0.019)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Shock Correlation Matrix**

<table>
<thead>
<tr>
<th>$r^m$</th>
<th>$K_t$</th>
<th>$y_t$</th>
<th>$\Delta \psi_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r^m$</td>
<td>1.000</td>
<td>-0.133</td>
<td>-0.009</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.051)</td>
<td>(0.047)</td>
</tr>
<tr>
<td>$K_t$</td>
<td>-0.133</td>
<td>1.000</td>
<td>-0.198</td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.000)</td>
<td>(0.055)</td>
</tr>
<tr>
<td>$y_t$</td>
<td>-0.009</td>
<td>-0.198</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td>(0.055)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>$\Delta \psi_t$</td>
<td>0.357</td>
<td>0.155</td>
<td>0.047</td>
</tr>
<tr>
<td></td>
<td>(0.073)</td>
<td>(0.046)</td>
<td>(0.049)</td>
</tr>
</tbody>
</table>
Table 1.3: Net Return Variance Decomposition (Proportional Cost) – This table reports the properties of dividend news ($\eta_d$), proportional cost news ($\eta_K$), and discount rate news ($\eta_R$) implied by the VAR model of Table 1.2. The first panel reports the annualized covariance matrix of the news terms (multiplied by 100) on the left and the correlation matrix on the right. The second panel reports the correlation of shocks to individual state variables with the three news terms on the left and the functions that map the state-variable shocks to the news terms on the right. The functional mappings are normalized by multiplying the term by its respective state variable shock volatility and dividing by the volatility of unexpected contemporaneous returns. Panel C provides seven measures of liquidity’s contribution to portfolio risk. $R_d$, $R_K$, and $R_y$ are respectively the ratios of dividend news, proportional cost news, and net discount rate news to that of unexpected contemporaneous returns. $\bar{P}_d$, $\bar{P}_K$, and $\bar{P}_y$ are the ratio of proportional cost news volatility to that of unexpected contemporaneous proportional costs. $B_d$, $B_K$, and $B_y$ are the regression coefficients obtained by regressing, respectively, dividend news, proportional cost news, and net discount rate news on unexpected contemporaneous returns. Net returns, proportional costs, log dividend yield, and log turnover growth rates are represented by $r^m$, $K$, $y$, and $\Delta \psi$ respectively. Included in brackets are small-sample bias corrected estimates using the jackknife procedure detailed in the appendix and robust jackknife standard errors are in parentheses.

<table>
<thead>
<tr>
<th>Panel A: News Covariance and Correlation Matrices</th>
<th>100 * News Covariance</th>
<th>News Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\eta_d$</td>
<td>6.9023</td>
<td>-0.0414</td>
</tr>
<tr>
<td>$\eta_K$</td>
<td>(5.2314)</td>
<td>[0.0498]</td>
</tr>
<tr>
<td>$\eta_R$</td>
<td>-0.0414</td>
<td>0.0008</td>
</tr>
<tr>
<td>$\eta_d$</td>
<td>(0.0498)</td>
<td>[0.0004]</td>
</tr>
<tr>
<td>$\eta_K$</td>
<td>(6.7624)</td>
<td>(0.0617)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Correlations and Shock → News Mappings</th>
<th>Shock Correlations</th>
<th>Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r^m$</td>
<td>0.341</td>
<td>0.030</td>
</tr>
<tr>
<td>$K$</td>
<td>0.062</td>
<td>0.716</td>
</tr>
<tr>
<td>$\eta_K$</td>
<td>(0.064)</td>
<td>(0.784)</td>
</tr>
<tr>
<td>$y$</td>
<td>0.886</td>
<td>-0.807</td>
</tr>
<tr>
<td>$\Delta \psi$</td>
<td>0.156</td>
<td>0.215</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C: Measures of Liquidity’s Contribution to Portfolio Risk</th>
<th>$\bar{P}_d$</th>
<th>$\bar{P}_K$</th>
<th>$\bar{P}_y$</th>
<th>$B_d$</th>
<th>$B_K$</th>
<th>$B_y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\eta_d$</td>
<td>6.904</td>
<td>1.433</td>
<td>0.016</td>
<td>1.450</td>
<td>0.489</td>
<td>0.000</td>
</tr>
<tr>
<td>$\eta_K$</td>
<td>[5.206]</td>
<td>[1.042]</td>
<td>[0.012]</td>
<td>[1.071]</td>
<td>[0.638]</td>
<td>[-0.001]</td>
</tr>
<tr>
<td>$\eta_R$</td>
<td>(2.326)</td>
<td>(0.577)</td>
<td>(0.006)</td>
<td>(0.824)</td>
<td>(0.380)</td>
<td>(0.004)</td>
</tr>
</tbody>
</table>

44
Table 1.4: Net Return Variance Decomposition (Fixed Cost) – This table reports the properties of dividend news ($\tilde{\eta}_d$), fixed cost news ($\tilde{\eta}_c$), and discount rate news ($\tilde{\eta}_r$) implied by the VAR model of Table 1.2. The first panel reports the annualized covariance matrix of the news terms (multiplied by 100) on the left and the correlation matrix on the right. The second panel reports the correlation of shocks to individual state variables with the three new terms on the left and the functions that map the state-variable shocks to the news terms on the right. The functional mappings are normalized by multiplying the term by its respective state variable shock volatility and dividing by the volatility of unexpected contemporaneous returns. Panel C provides seven measures of liquidity’s contribution to portfolio risk. $\tilde{\mathcal{R}}_d$, $\tilde{\mathcal{R}}_K$, and $\tilde{\mathcal{R}}_f$ are respectively the ratios of dividend news, proportional cost news, and net discount rate news to that of unexpected contemporaneous returns. $\tilde{\mathcal{P}}_d$ is the ratio of proportional cost news volatility to that of unexpected contemporaneous proportional costs. $\tilde{\mathcal{B}}_d$, $\tilde{\mathcal{B}}_c$, and $\tilde{\mathcal{B}}_f$ are the regression coefficients obtained by regressing, respectively, dividend news, proportional cost news, and net discount rate news on unexpected contemporaneous returns. Net returns, proportional costs, log dividend yield, and log turnover growth rates are represented by $r^{m}$, $K$, $y$, and $\Delta\psi$ respectively. Included in brackets are small-sample bias corrected estimates using the jackknife procedure detailed in the appendix and robust jackknife standard errors are in parentheses.

### Panel A: News Covariance and Correlation Matrices

<table>
<thead>
<tr>
<th>100 * News Covariance</th>
<th>News Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tilde{\eta}_d$</td>
<td>10.6382</td>
</tr>
<tr>
<td></td>
<td>(8.5075)</td>
</tr>
<tr>
<td>$\tilde{\eta}_c$</td>
<td>2.0177</td>
</tr>
<tr>
<td></td>
<td>[0.6623]</td>
</tr>
<tr>
<td></td>
<td>(1.6001)</td>
</tr>
<tr>
<td>$\tilde{\eta}_f$</td>
<td>6.5818</td>
</tr>
<tr>
<td></td>
<td>[0.0886]</td>
</tr>
<tr>
<td></td>
<td>(8.3934)</td>
</tr>
</tbody>
</table>

### Panel B: Correlations and Shock -> News Mappings

<table>
<thead>
<tr>
<th>Shock</th>
<th>$r^{m}$</th>
<th>$K$</th>
<th>$y$</th>
<th>$\Delta\psi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tilde{\eta}_d$</td>
<td>0.341</td>
<td>0.351</td>
<td>-0.353</td>
<td>0.756</td>
</tr>
<tr>
<td></td>
<td>[0.388]</td>
<td>[0.400]</td>
<td>[-0.370]</td>
<td>[0.914]</td>
</tr>
<tr>
<td>$\tilde{\eta}_c$</td>
<td>0.062</td>
<td>0.096</td>
<td>0.146</td>
<td>0.573</td>
</tr>
<tr>
<td></td>
<td>[0.094]</td>
<td>[0.103]</td>
<td>[0.145]</td>
<td>[0.454]</td>
</tr>
<tr>
<td>$\tilde{\eta}_f$</td>
<td>0.886</td>
<td>0.871</td>
<td>0.891</td>
<td>1.705</td>
</tr>
<tr>
<td></td>
<td>[0.973]</td>
<td>[0.957]</td>
<td>[0.904]</td>
<td>[1.295]</td>
</tr>
</tbody>
</table>

### Panel C: Measures of Liquidity’s Contribution to Portfolio Risk

<table>
<thead>
<tr>
<th>$\tilde{\mathcal{R}}_d$</th>
<th>$\tilde{\mathcal{R}}_K$</th>
<th>$\tilde{\mathcal{R}}_f$</th>
<th>$\tilde{\mathcal{B}}_d$</th>
<th>$\tilde{\mathcal{B}}_c$</th>
<th>$\tilde{\mathcal{B}}_f$</th>
</tr>
</thead>
<tbody>
<tr>
<td>144.464</td>
<td>1.779</td>
<td>0.338</td>
<td>1.450</td>
<td>0.607</td>
<td>0.118</td>
</tr>
<tr>
<td>[105.270]</td>
<td>[1.294]</td>
<td>[0.246]</td>
<td>[1.071]</td>
<td>[0.792]</td>
<td>[0.153]</td>
</tr>
<tr>
<td>(57.492)</td>
<td>(0.716)</td>
<td>(0.134)</td>
<td>(0.824)</td>
<td>(0.472)</td>
<td>(0.090)</td>
</tr>
</tbody>
</table>
Table 1.5: Size Sorted Portfolios – This table reports statistics for quintile-ranked portfolios. At the end of each year, beginning in 1963 and ending 2001, eligible stocks are sorted into five portfolios. The breakpoints are based on all eligible stocks so each portfolio has approximately the same number of stocks at the time of formation. The portfolio characteristics for the 12 post-ranking months are joined across years to form a single time series for each quintile. Panel A reports the annualized sample statistics including the time series average of log net returns, variance of log net returns, variance of the unexpected component of log net returns, market capitalization, yield, proportional illiquidity level, and turnover. Market capitalization is reported in billions of dollars. Everything else is reported in percent per year. Panel B presents the statistics associated with the proportional cost decomposition. \( B_d \), \( B_K \), and \( B_c \) are, respectively, the regression coefficients of dividend news, proportion cost news, and net discount rate news on unexpected contemporaneous returns. \( R_K \) and \( \tilde{P} \) are the volatility of proportional cost news normalized by the volatility of, respectively, unexpected returns and contemporaneous proportional costs. Panel C presents the statistics associated with the fixed cost decomposition. \( B_d \), \( B_c \), and \( B_p \) are, respectively, the regression coefficients of dividend news, fixed cost news, and net discount rate news on unexpected contemporaneous returns. \( R_d \) and \( \tilde{P} \) are, respectively, the relative volatilities of dividend news and proportional cost news normalized by fixed cost news. All estimates reported in Panels B through D are bias-corrected using the small-sample jackknife bias correction procedure detailed in the Appendix. Jackknife standard errors are presented in parenthesis.

<table>
<thead>
<tr>
<th>Sort by Market Capitalization</th>
<th>Decile</th>
<th>Low</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>High</th>
<th>Avg</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel A: Sample Statistics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( E[\tilde{r}_i] )</td>
<td>11.130</td>
<td>11.2512</td>
<td>11.8543</td>
<td>11.9114</td>
<td>10.6104</td>
<td>12.3193</td>
<td></td>
</tr>
<tr>
<td>( \sigma (\tilde{r}) )</td>
<td>21.1850</td>
<td>20.9167</td>
<td>19.9318</td>
<td>18.0530</td>
<td>15.7682</td>
<td>18.7700</td>
<td></td>
</tr>
<tr>
<td>( \sigma (\nu) )</td>
<td>20.3804</td>
<td>20.3926</td>
<td>19.6474</td>
<td>17.9668</td>
<td>15.8289</td>
<td>18.4839</td>
<td></td>
</tr>
<tr>
<td>Market Capitalization</td>
<td>0.0330</td>
<td>0.1136</td>
<td>0.2982</td>
<td>0.8233</td>
<td>6.2615</td>
<td>1.5852</td>
<td></td>
</tr>
<tr>
<td>Yield</td>
<td>0.0419</td>
<td>2.4514</td>
<td>2.8127</td>
<td>3.2181</td>
<td>3.5776</td>
<td>2.8278</td>
<td></td>
</tr>
<tr>
<td>Cost (Chalmers and Kadlec)</td>
<td>1.2332</td>
<td>0.6677</td>
<td>0.4164</td>
<td>0.2648</td>
<td>0.1343</td>
<td>0.5005</td>
<td></td>
</tr>
<tr>
<td>Cost (Acharya and Pedersen)</td>
<td>2.1470</td>
<td>0.7040</td>
<td>0.4357</td>
<td>0.2901</td>
<td>0.1887</td>
<td>0.6272</td>
<td></td>
</tr>
<tr>
<td>Turnover</td>
<td>48.1750</td>
<td>55.9939</td>
<td>60.0665</td>
<td>62.0151</td>
<td>54.9749</td>
<td>58.0115</td>
<td></td>
</tr>
</tbody>
</table>

| **Panel B: Proportional Cost Decomposition** |        |     |   |   |   |      |     |
| \( B_d \)                     | 0.6708 | 0.7013 | 0.8437 | 0.8969 | 0.8258 | 0.6382 |
| \( B_K \)                     | -0.0242 | -0.0054 | -0.002 | 0.0004 | 0.0002 | -0.0007 |
| \( B_c \)                     | -0.3050 | -0.2932 | -0.1561 | -0.1035 | -0.1744 | -0.3611 |
| \( R_K \)                     | 0.0560 | 0.0124 | 0.0071 | 0.0044 | 0.0023 | 0.0121 |
| \( \tilde{P} \)                | 10.8337 | 4.3580 | 2.5245 | 2.5249 | 2.3705 | 5.2063 |
|                             | (3.7956) | (1.8547) | (0.3306) | (1.1587) | (0.8171) | (2.3258) |

| **Panel C: Fixed Cost Decomposition** |        |     |   |   |   |      |     |
| \( B_d \)                     | 1.0582 | 0.9114 | 1.0180 | 1.0636 | 0.9245 | 0.7837 |
| \( B_c \)                     | 0.3632 | 0.2046 | 0.1741 | 0.1400 | 0.0990 | 0.1448 |
| \( B_p \)                     | -0.3050 | -0.2932 | -0.1561 | -0.1035 | -0.1744 | -0.3611 |
| \( R_c \)                     | 0.5984 | 0.3961 | 0.2510 | 0.1617 | 0.1236 | 0.2324 |
| \( \tilde{P} \)                | 117.0275 | 138.9953 | 88.5751 | 90.3083 | 123.5354 | 99.3216 |
|                             | (20.2569) | (44.8135) | (32.1982) | (41.9418) | (42.5141) | (54.1943) |

| **Panel D: Decomposing Fixed Cost News** |        |     |   |   |   |      |     |
| \( B_d^* \)                | 1.0519 | 1.0003 | 1.0046 | 1.0098 | 1.0076 | 1.0259 |
| \( B_K^* \)                | -0.0519 | -0.0003 | -0.0046 | -0.0098 | -0.0076 | -0.0259 |
| \( R_d^* \)                | 1.0546 | 1.0007 | 1.0050 | 1.0100 | 1.0078 | 1.0267 |
| \( R_K^* \)                | 0.0931 | 0.0293 | 0.0261 | 0.0253 | 0.0182 | 0.0472 |

46
Table 1.6: Illiquidity Sorted Portfolios – This table reports statistics for quintile-ranked portfolios. At the end of each year, beginning in 1963 and ending 2001, eligible stocks are sorted into five portfolios. The breakpoints are based on all eligible stocks so each portfolio has approximately the same number of stocks at the time of formation. The portfolio characteristics for the 12 post-ranking months are joined across years to form a single time series for each quintile. Panel A reports the annualized sample statistics including the time series average of log net returns, variance of log net returns, variance of the unexpected component of log net returns, market capitalization, yield, proportional illiquidity level, and turnover. Market capitalization is reported in billions of dollars. Everything else is reported in percent per year. Panel B presents the statistics associated with the proportional cost decomposition. \( B_d \), \( B_K \), and \( B_p \) are, respectively, the regression coefficients of dividend news, proportion cost news, and net discount rate news on unexpected contemporaneous returns. \( R_F \) and \( p_F \) are the volatility of proportional cost news normalized by the volatility of, respectively, unexpected returns and contemporaneous proportional costs. Panel C presents the statistics associated with the fixed cost decomposition. \( B_d^* \), \( B_K^* \), and \( B_p^* \) are, respectively, the regression coefficients of dividend news, fixed cost news, and net discount rate news on unexpected contemporaneous returns. \( R_F^* \) and \( p_F^* \) are the volatility of proportional cost news normalized by the volatility of, respectively, unexpected returns and contemporaneous proportional costs. Panel D reports the statistics associated with the decomposition of fixed cost news into dividend news and proportional cost news. \( B_d^* \) and \( B_K^* \) are, respectively, the regression coefficients of dividend news and proportional cost news on fixed cost news. \( R_F^* \) and \( p_F^* \) are, respectively, the relative volatilities of dividend news and proportional cost news normalized by fixed cost news. All estimates reported in Panels B through D are bias-corrected using the small-sample jackknife bias correction procedure detailed in the Appendix. Jackknife standard errors are presented in parenthesis.

<table>
<thead>
<tr>
<th>Sort by Acharya and Pedersen (2005) Illiquidity Levels</th>
<th>Decile</th>
<th>Low</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>High</th>
<th>Avg</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Sample Statistics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( E(r_i^*) )</td>
<td>10.3826</td>
<td>10.8234</td>
<td>11.2022</td>
<td>11.6542</td>
<td>12.1418</td>
<td>12.3193</td>
<td></td>
</tr>
<tr>
<td>( \sigma(e_i) )</td>
<td>0.2685</td>
<td>0.1537</td>
<td>0.2249</td>
<td>0.2337</td>
<td>0.2066</td>
<td>0.3796</td>
<td></td>
</tr>
<tr>
<td>( \sigma(e_p) )</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0014</td>
<td>0.0030</td>
<td>0.0155</td>
<td>0.0040</td>
<td></td>
</tr>
<tr>
<td>Yield</td>
<td>6.1660</td>
<td>8.7097</td>
<td>0.3363</td>
<td>0.1443</td>
<td>0.0487</td>
<td>1.5852</td>
<td></td>
</tr>
<tr>
<td>Market Capitalization</td>
<td>3.5153</td>
<td>3.1130</td>
<td>2.7737</td>
<td>2.5061</td>
<td>2.2027</td>
<td>2.8278</td>
<td></td>
</tr>
<tr>
<td>Cost (Chalmers and Kadlec)</td>
<td>0.1263</td>
<td>0.2503</td>
<td>0.4040</td>
<td>0.6715</td>
<td>1.2620</td>
<td>0.5050</td>
<td></td>
</tr>
<tr>
<td>Cost (Acharya and Pedersen)</td>
<td>0.1820</td>
<td>0.2741</td>
<td>0.4218</td>
<td>0.7029</td>
<td>2.1840</td>
<td>0.6272</td>
<td></td>
</tr>
<tr>
<td>Turnover</td>
<td>65.9823</td>
<td>67.6265</td>
<td>59.8922</td>
<td>48.7651</td>
<td>38.6905</td>
<td>58.0115</td>
<td></td>
</tr>
<tr>
<td><strong>Panel B: Proportional Cost Decomposition</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( B_d )</td>
<td>0.0002</td>
<td>-0.0005</td>
<td>-0.0005</td>
<td>-0.0045</td>
<td>-0.2223</td>
<td>-0.0007</td>
<td></td>
</tr>
<tr>
<td>( B_K )</td>
<td>0.0017</td>
<td>0.0043</td>
<td>0.0067</td>
<td>0.0131</td>
<td>0.0670</td>
<td>0.0121</td>
<td></td>
</tr>
<tr>
<td>( B_p )</td>
<td>2.0544</td>
<td>2.7244</td>
<td>2.3204</td>
<td>3.9285</td>
<td>12.8278</td>
<td>5.2063</td>
<td></td>
</tr>
<tr>
<td>( R_F )</td>
<td>5.4839</td>
<td>3.193</td>
<td>0.3953</td>
<td>1.2278</td>
<td>3.8282</td>
<td>2.3258</td>
<td></td>
</tr>
<tr>
<td><strong>Panel C: Fixed Cost Decomposition</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( B_d )</td>
<td>0.8999</td>
<td>1.1330</td>
<td>1.0949</td>
<td>1.1419</td>
<td>0.9991</td>
<td>0.7837</td>
<td></td>
</tr>
<tr>
<td>( B_K )</td>
<td>0.0961</td>
<td>0.1528</td>
<td>0.1859</td>
<td>0.2569</td>
<td>0.3374</td>
<td>0.1448</td>
<td></td>
</tr>
<tr>
<td>( B_p )</td>
<td>-0.1962</td>
<td>-0.0199</td>
<td>-0.0911</td>
<td>-0.1150</td>
<td>-0.3383</td>
<td>-0.3611</td>
<td></td>
</tr>
<tr>
<td>( R_F )</td>
<td>1.126</td>
<td>0.1850</td>
<td>0.2376</td>
<td>0.3915</td>
<td>0.5771</td>
<td>0.2324</td>
<td></td>
</tr>
<tr>
<td><strong>Panel D: Decomposing Fixed Cost News</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( B_d^* )</td>
<td>1.0079</td>
<td>1.0085</td>
<td>1.0026</td>
<td>1.0016</td>
<td>1.0728</td>
<td>1.0259</td>
<td></td>
</tr>
<tr>
<td>( B_K^* )</td>
<td>-0.0079</td>
<td>-0.0085</td>
<td>-0.0026</td>
<td>-0.0016</td>
<td>-0.0728</td>
<td>-0.0259</td>
<td></td>
</tr>
<tr>
<td>( R_F^* )</td>
<td>1.0080</td>
<td>1.0087</td>
<td>1.0030</td>
<td>1.0021</td>
<td>1.0764</td>
<td>1.0267</td>
<td></td>
</tr>
<tr>
<td>( R_K^* )</td>
<td>0.0145</td>
<td>0.0022</td>
<td>0.0259</td>
<td>0.0313</td>
<td>0.1149</td>
<td>0.0472</td>
<td></td>
</tr>
</tbody>
</table>
Table 1.7: Turnover Sorted Portfolios – This table reports statistics for quintile-ranked portfolios. At the end of each year, beginning in 1963 and ending 2001, eligible stocks are sorted into five portfolios. The breakpoints are based on all eligible stocks so each portfolio has approximately the same number of stocks at the time of formation. The portfolio characteristics for the 12 post-ranking months are joined across years to form a single time series for each quintile. Panel A reports the annualized sample statistics including the time series average of log net returns, variance of log net returns, variance of the unexpected component of log net returns, market capitalization, yield, proportional illiquidity level, and turnover. Market capitalization is reported in billions of dollars. Everything else is reported in percent per year. Panel B presents the statistics associated with the proportional cost decomposition. $\beta_d$, $\beta_K$, and $\beta_c$ are, respectively, the regression coefficients of dividend news, proportion cost news, and net discount rate news on unexpected contemporaneous returns. $\beta^K_{d}$ and $\beta^K_{c}$ are the volatility of proportional cost news normalized by the volatility of, respectively, unexpected returns and contemporaneous proportional costs. Panel C presents the statistics associated with the fixed cost decomposition. $\beta_d$, $\beta_c$, and $\beta_p$ are, respectively, the regression coefficients of dividend news, fixed cost news, and net discount rate news on unexpected contemporaneous returns. $\beta^K_{d}$ and $\beta^K_{c}$ are, respectively, the relative volatilities of dividend news and proportional cost news normalized by fixed cost news. All estimates reported in Panels B through D are bias-corrected using the small-sample jackknife bias correction procedure detailed in the Appendix. Jackknife standard errors are presented in parenthesis.

<table>
<thead>
<tr>
<th>Sort by Turnover</th>
<th>Decile</th>
<th>Low</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>High</th>
<th>Avg</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Sample Statistics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma(\gamma)$</td>
<td>14.0390</td>
<td>16.0207</td>
<td>18.2855</td>
<td>20.9888</td>
<td>25.1810</td>
<td>18.7700</td>
<td></td>
</tr>
<tr>
<td>$\sigma(\varphi)$</td>
<td>13.7340</td>
<td>15.7453</td>
<td>18.0605</td>
<td>20.7142</td>
<td>24.9622</td>
<td>18.4839</td>
<td></td>
</tr>
<tr>
<td>Market Capitalization</td>
<td>0.7277</td>
<td>2.0611</td>
<td>2.0016</td>
<td>1.7058</td>
<td>1.1344</td>
<td>1.5852</td>
<td></td>
</tr>
<tr>
<td>Yield</td>
<td>3.6212</td>
<td>3.4767</td>
<td>3.0380</td>
<td>2.4161</td>
<td>1.5566</td>
<td>2.8278</td>
<td></td>
</tr>
<tr>
<td>Cost (Chalmers and Kadlec)</td>
<td>0.7988</td>
<td>0.5062</td>
<td>0.4868</td>
<td>0.4505</td>
<td>0.4041</td>
<td>0.5005</td>
<td></td>
</tr>
<tr>
<td>Cost (Acharya and Pedersen)</td>
<td>1.2807</td>
<td>0.7292</td>
<td>0.6087</td>
<td>0.5698</td>
<td>0.5540</td>
<td>0.6272</td>
<td></td>
</tr>
<tr>
<td>Turnover</td>
<td>17.9624</td>
<td>34.2872</td>
<td>48.5751</td>
<td>68.0489</td>
<td>113.0961</td>
<td>58.0115</td>
<td></td>
</tr>
<tr>
<td><strong>Panel B: Proportional Cost Decomposition</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta^K_{d}$</td>
<td>0.6775</td>
<td>0.8471</td>
<td>0.7513</td>
<td>0.7700</td>
<td>0.7651</td>
<td>0.6382</td>
<td></td>
</tr>
<tr>
<td>$\beta^K_{c}$</td>
<td>0.0190</td>
<td>0.0065</td>
<td>0.0059</td>
<td>0.0038</td>
<td>0.0010</td>
<td>0.0007</td>
<td></td>
</tr>
<tr>
<td>$\beta^K_{p}$</td>
<td>0.3035</td>
<td>0.1464</td>
<td>0.2429</td>
<td>0.2262</td>
<td>0.2338</td>
<td>0.3611</td>
<td></td>
</tr>
<tr>
<td>$\beta^K_{d}$</td>
<td>0.1141</td>
<td>0.0812</td>
<td>0.0139</td>
<td>0.0111</td>
<td>0.0089</td>
<td>0.0121</td>
<td></td>
</tr>
<tr>
<td>$\beta^K_{c}$</td>
<td>0.0881</td>
<td>0.0047</td>
<td>0.0033</td>
<td>0.0035</td>
<td>0.0017</td>
<td>0.0057</td>
<td></td>
</tr>
<tr>
<td>$\beta^K_{p}$</td>
<td>27.8563</td>
<td>6.4179</td>
<td>5.5295</td>
<td>4.6044</td>
<td>3.8382</td>
<td>5.2063</td>
<td></td>
</tr>
<tr>
<td>Turnover</td>
<td>21.6001</td>
<td>1.5782</td>
<td>1.3779</td>
<td>1.5126</td>
<td>0.7629</td>
<td>2.3258</td>
<td></td>
</tr>
<tr>
<td><strong>Panel C: Fixed Cost Decomposition</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta^K_{d}$</td>
<td>0.8566</td>
<td>1.0281</td>
<td>0.9125</td>
<td>0.9515</td>
<td>0.9778</td>
<td>0.7837</td>
<td></td>
</tr>
<tr>
<td>$\beta^K_{c}$</td>
<td>0.1601</td>
<td>0.1745</td>
<td>0.1554</td>
<td>0.1777</td>
<td>0.2116</td>
<td>0.1448</td>
<td></td>
</tr>
<tr>
<td>$\beta^K_{p}$</td>
<td>0.3035</td>
<td>0.1464</td>
<td>0.2429</td>
<td>0.2262</td>
<td>0.2338</td>
<td>0.3611</td>
<td></td>
</tr>
<tr>
<td>$\beta^K_{d}$</td>
<td>0.2093</td>
<td>0.2370</td>
<td>0.2389</td>
<td>0.2610</td>
<td>0.3108</td>
<td>0.2324</td>
<td></td>
</tr>
<tr>
<td>$\beta^K_{c}$</td>
<td>51.0513</td>
<td>82.9078</td>
<td>94.8345</td>
<td>108.1766</td>
<td>131.9727</td>
<td>99.3216</td>
<td></td>
</tr>
<tr>
<td>Turnover</td>
<td>9.5366</td>
<td>27.1984</td>
<td>21.8071</td>
<td>32.7916</td>
<td>47.1805</td>
<td>54.1943</td>
<td></td>
</tr>
<tr>
<td><strong>Panel D: Decomposing Fixed Cost News</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta^K_{d}$</td>
<td>1.1549</td>
<td>1.0324</td>
<td>1.0152</td>
<td>1.0116</td>
<td>1.0045</td>
<td>1.0259</td>
<td></td>
</tr>
<tr>
<td>$\beta^K_{c}$</td>
<td>0.1549</td>
<td>0.0324</td>
<td>0.0152</td>
<td>0.0116</td>
<td>0.0045</td>
<td>0.0259</td>
<td></td>
</tr>
<tr>
<td>$\beta^K_{d}$</td>
<td>1.2662</td>
<td>1.0344</td>
<td>1.0166</td>
<td>1.0123</td>
<td>1.0048</td>
<td>1.0267</td>
<td></td>
</tr>
<tr>
<td>$\beta^K_{c}$</td>
<td>0.5531</td>
<td>0.0734</td>
<td>0.0563</td>
<td>0.0403</td>
<td>0.0262</td>
<td>0.0472</td>
<td></td>
</tr>
</tbody>
</table>

48
value-weighted portfolios and an equally-weighted market portfolio. The beta is calculated using the decomposed returns of the portfolios and an equally-weighted market portfolio. $\beta^*$ is the portfolio's liquidity-adjusted CAPM beta, $\beta_{r,c}$ is the portfolio's non-liquidity news sensitivity to the market's liquidity news, $\beta_{c,r}$ is the portfolio's non-liquidity news sensitivity to the market's liquidity news, $\beta_{c,r}$ is the portfolio's liquidity news sensitivity to the market's liquidity news, and $\beta_{c,c}$ is the portfolio's liquidity news sensitivity to the market's liquidity news. The reported betas are small-sample bias corrected using the jackknife procedure. Robust standard errors are reported in parenthesis. In addition, for each portfolio, the average market capitalization, annualized time-series expectation and volatility of proportional transactions costs and net excess returns, turnover, and yield are reported.

### Table 1.8: Betas: Contemporaneous Decomposition

This table reports the properties of the odd-numbered value-weighted portfolios obtained by sorting eligible stocks each year, beginning in 1963 and ending in 1999, according to their annual Amihud (2002) illiquidity measures. The breakpoints are based on eligible stocks so each portfolio has approximately the same number of stocks at formation. The betas are calculated using the decomposed returns of the value-weighted portfolios and an equally-weighted market portfolio. $\beta^*$ is the portfolio's liquidity-adjusted CAPM beta, $\beta_{r,c}$ is the portfolio's non-liquidity news sensitivity to the market's liquidity news, $\beta_{c,r}$ is the portfolio's non-liquidity news sensitivity to the market's liquidity news, $\beta_{c,c}$ is the portfolio's liquidity news sensitivity to the market's liquidity news. The reported betas are small-sample bias corrected using the jackknife procedure. Robust standard errors are reported in parenthesis. In addition, for each portfolio, the average market capitalization, annualized time-series expectation and volatility of proportional transactions costs and net excess returns, turnover, and yield are reported.

<table>
<thead>
<tr>
<th></th>
<th>$\beta^*$</th>
<th>$\beta_{r,c}$</th>
<th>$\beta_{c,r}$</th>
<th>$\beta_{c,c}$</th>
<th>$E(K)$</th>
<th>$\sigma(K)$</th>
<th>$E(\tau)$</th>
<th>$\sigma(\tau)$</th>
<th>$\psi$</th>
<th>$y$</th>
<th>mktcap</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>53.05</td>
<td>52.96</td>
<td>-0.09</td>
<td>-0.01</td>
<td>-0.00</td>
<td>0.16</td>
<td>0.02</td>
<td>4.80</td>
<td>14.59</td>
<td>39.37</td>
<td>3.64</td>
</tr>
<tr>
<td>2</td>
<td>68.97</td>
<td>68.88</td>
<td>-0.09</td>
<td>0.00</td>
<td>0.00</td>
<td>0.17</td>
<td>0.02</td>
<td>4.33</td>
<td>16.02</td>
<td>51.14</td>
<td>3.72</td>
</tr>
<tr>
<td>3</td>
<td>75.61</td>
<td>75.48</td>
<td>-0.12</td>
<td>-0.00</td>
<td>0.01</td>
<td>0.21</td>
<td>0.05</td>
<td>5.51</td>
<td>16.76</td>
<td>52.39</td>
<td>3.74</td>
</tr>
<tr>
<td>4</td>
<td>79.54</td>
<td>79.44</td>
<td>-0.10</td>
<td>0.01</td>
<td>0.01</td>
<td>0.22</td>
<td>0.05</td>
<td>6.16</td>
<td>17.02</td>
<td>52.75</td>
<td>3.26</td>
</tr>
<tr>
<td>5</td>
<td>85.70</td>
<td>85.56</td>
<td>-0.15</td>
<td>0.01</td>
<td>0.01</td>
<td>0.29</td>
<td>0.07</td>
<td>5.95</td>
<td>18.09</td>
<td>48.68</td>
<td>3.23</td>
</tr>
<tr>
<td>6</td>
<td>90.68</td>
<td>90.50</td>
<td>-0.14</td>
<td>-0.03</td>
<td>0.01</td>
<td>0.34</td>
<td>0.09</td>
<td>5.37</td>
<td>18.50</td>
<td>49.12</td>
<td>3.15</td>
</tr>
<tr>
<td>7</td>
<td>91.81</td>
<td>91.67</td>
<td>-0.17</td>
<td>0.04</td>
<td>0.02</td>
<td>0.41</td>
<td>0.11</td>
<td>7.00</td>
<td>18.57</td>
<td>45.66</td>
<td>2.95</td>
</tr>
<tr>
<td>8</td>
<td>92.75</td>
<td>92.64</td>
<td>-0.14</td>
<td>0.05</td>
<td>0.02</td>
<td>0.51</td>
<td>0.15</td>
<td>7.79</td>
<td>19.34</td>
<td>41.90</td>
<td>2.80</td>
</tr>
<tr>
<td>9</td>
<td>94.33</td>
<td>94.13</td>
<td>-0.14</td>
<td>-0.04</td>
<td>0.02</td>
<td>0.55</td>
<td>0.12</td>
<td>7.63</td>
<td>19.38</td>
<td>39.09</td>
<td>2.72</td>
</tr>
<tr>
<td>10</td>
<td>94.26</td>
<td>93.96</td>
<td>-0.14</td>
<td>-0.14</td>
<td>0.02</td>
<td>0.69</td>
<td>0.15</td>
<td>7.34</td>
<td>19.48</td>
<td>36.64</td>
<td>2.63</td>
</tr>
<tr>
<td>11</td>
<td>98.00</td>
<td>97.65</td>
<td>-0.10</td>
<td>-0.23</td>
<td>0.02</td>
<td>0.85</td>
<td>0.14</td>
<td>9.31</td>
<td>20.30</td>
<td>34.60</td>
<td>2.54</td>
</tr>
<tr>
<td>12</td>
<td>104.78</td>
<td>104.01</td>
<td>-0.10</td>
<td>-0.63</td>
<td>0.03</td>
<td>1.38</td>
<td>0.21</td>
<td>9.21</td>
<td>21.23</td>
<td>37.93</td>
<td>2.19</td>
</tr>
<tr>
<td>13</td>
<td>101.27</td>
<td>99.21</td>
<td>-0.07</td>
<td>-1.92</td>
<td>0.06</td>
<td>3.13</td>
<td>0.52</td>
<td>10.36</td>
<td>21.33</td>
<td>34.17</td>
<td>2.10</td>
</tr>
</tbody>
</table>
Table 1.9: Betas: Proportional Cost Decomposition—This table reports the properties of the odd-numbered value-weighted portfolios obtained by sorting eligible stocks each year, beginning in 1963 and ending in 1999, according to their annual Amihud (2002) illiquidity measures. The breakpoints are based on eligible stocks so each portfolio has approximately the same number of stocks at formation. The betas are calculated using the decomposed returns of the value-weighted portfolios and an equally-weighted market portfolio. $\beta^*$ is the portfolio’s liquidity-adjusted CAPM beta, $\beta_{r,c}$ is the portfolio’s non-liquidity news sensitivity to the market’s non-liquidity news, $\beta_{r,c}$ is the portfolio’s non-liquidity news sensitivity to the market’s liquidity news, $\beta_{c,r}$ is the portfolio’s liquidity news sensitivity to the market’s non-liquidity news, and $\beta_{c,c}$ is the portfolio’s liquidity news sensitivity to the market’s liquidity news. The reported betas are small-sample bias corrected using the jackknife procedure. Robust standard errors are reported in parenthesis. In addition, for each portfolio, the average market capitalization, annualized time-series expectation and volatility of proportional transactions costs and net excess returns, turnover, and yield are reported.

<table>
<thead>
<tr>
<th>Illiquidity-Ranked Portfolios: Proportional Cost Decomposition</th>
<th>$\beta^*$</th>
<th>$\beta_{r,r}$</th>
<th>$\beta_{r,c}$</th>
<th>$\beta_{c,r}$</th>
<th>$\beta_{c,c}$</th>
<th>$\psi(K)$</th>
<th>$\sigma(K)$</th>
<th>$\psi(\ell)$</th>
<th>$\psi(y)$</th>
<th>$\psi(\text{mktcap})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>64.07</td>
<td>64.00</td>
<td>−0.09</td>
<td>0.02</td>
<td>0.00</td>
<td>0.16</td>
<td>0.02</td>
<td>4.80</td>
<td>14.59</td>
<td>39.37</td>
</tr>
<tr>
<td>2</td>
<td>77.89</td>
<td>77.79</td>
<td>−0.11</td>
<td>0.01</td>
<td>0.00</td>
<td>0.17</td>
<td>0.02</td>
<td>4.33</td>
<td>16.02</td>
<td>51.14</td>
</tr>
<tr>
<td>3</td>
<td>82.43</td>
<td>82.39</td>
<td>−0.13</td>
<td>0.10</td>
<td>0.00</td>
<td>0.21</td>
<td>0.05</td>
<td>5.51</td>
<td>16.76</td>
<td>52.39</td>
</tr>
<tr>
<td>4</td>
<td>85.14</td>
<td>85.04</td>
<td>−0.09</td>
<td>−0.02</td>
<td>0.00</td>
<td>0.22</td>
<td>0.05</td>
<td>6.16</td>
<td>17.02</td>
<td>52.75</td>
</tr>
<tr>
<td>5</td>
<td>91.34</td>
<td>91.17</td>
<td>−0.14</td>
<td>−0.03</td>
<td>0.00</td>
<td>0.29</td>
<td>0.07</td>
<td>5.95</td>
<td>18.09</td>
<td>48.68</td>
</tr>
<tr>
<td>6</td>
<td>94.39</td>
<td>94.27</td>
<td>−0.13</td>
<td>0.01</td>
<td>0.01</td>
<td>0.34</td>
<td>0.09</td>
<td>5.37</td>
<td>18.50</td>
<td>49.12</td>
</tr>
<tr>
<td>7</td>
<td>94.79</td>
<td>94.84</td>
<td>−0.15</td>
<td>0.21</td>
<td>0.01</td>
<td>0.41</td>
<td>0.11</td>
<td>7.00</td>
<td>18.57</td>
<td>45.66</td>
</tr>
<tr>
<td>8</td>
<td>97.71</td>
<td>97.74</td>
<td>−0.12</td>
<td>0.16</td>
<td>0.01</td>
<td>0.51</td>
<td>0.15</td>
<td>7.79</td>
<td>19.34</td>
<td>41.90</td>
</tr>
<tr>
<td>9</td>
<td>97.39</td>
<td>97.25</td>
<td>−0.11</td>
<td>−0.02</td>
<td>0.01</td>
<td>0.55</td>
<td>0.12</td>
<td>7.63</td>
<td>19.38</td>
<td>39.09</td>
</tr>
<tr>
<td>10</td>
<td>95.80</td>
<td>95.63</td>
<td>−0.10</td>
<td>−0.06</td>
<td>0.01</td>
<td>0.69</td>
<td>0.15</td>
<td>7.34</td>
<td>19.48</td>
<td>36.64</td>
</tr>
<tr>
<td>11</td>
<td>98.49</td>
<td>97.47</td>
<td>−0.03</td>
<td>−0.97</td>
<td>0.01</td>
<td>0.85</td>
<td>0.14</td>
<td>9.31</td>
<td>20.30</td>
<td>34.60</td>
</tr>
<tr>
<td>12</td>
<td>97.67</td>
<td>95.67</td>
<td>−0.05</td>
<td>−1.93</td>
<td>0.02</td>
<td>1.38</td>
<td>0.21</td>
<td>9.21</td>
<td>21.23</td>
<td>37.93</td>
</tr>
<tr>
<td>13</td>
<td>95.24</td>
<td>89.68</td>
<td>0.06</td>
<td>−5.59</td>
<td>0.04</td>
<td>3.13</td>
<td>0.52</td>
<td>10.36</td>
<td>21.33</td>
<td>34.17</td>
</tr>
</tbody>
</table>
Table 1.10: Betas: Fixed Cost Decomposition – This table reports the properties of the odd-numbered value-weighted portfolios obtained by sorting eligible stocks each year, beginning in 1963 and ending in 1999, according to their annual Amihud (2002) illiquidity measures. The breakpoints are based on eligible stocks so each portfolio has approximately the same number of stocks at formation. The betas are calculated using the decomposed returns of the value-weighted portfolios and an equally-weighted market portfolio. $\beta^*$ is the portfolio’s liquidity-adjusted CAPM beta, $\beta_{r,c}$ is the portfolio’s non-liquidity news sensitivity to the market’s non-liquidity news, $\beta_{c,r}$ is the portfolio’s non-liquidity news sensitivity to the market’s liquidity news, $\beta_{r,c}$ is the portfolio’s liquidity news sensitivity to the market’s non-liquidity news, and $\beta_{c,r}$ is the portfolio’s liquidity news sensitivity to the market’s liquidity news. The reported betas are small-sample bias corrected using the jackknife procedure. Robust standard errors are reported in parenthesis. In addition, for each portfolio, the average market capitalization, annualized time-series expectation and volatility of proportional transactions costs and net excess returns, turnover, and yield are reported.

| Illiquidity-Ranked Portfolios: Fixed Cost Decomposition |
|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
|                | $\beta^*$     | $\beta_{r,c}$ | $\beta_{c,r}$ | $\beta_{r,c}$ | $E(K)$         | $\sigma(K)$    | $E(t)$         | $\sigma(t)$    | $\psi$         | $y$            | mktcap         |
| 1              | 64.07 (2.61)  | 79.85 (5.66)  | 9.92 (4.75)   | 7.04 (0.93)   | 1.17 (0.31)    | 0.16 (0.02)    | 4.80 (14.59)   | 39.37 (37.36)  | 3.63 (14.41)   |
| 3              | 77.89 (2.23)  | 98.85 (6.00)  | 13.11 (5.32)  | 9.41 (1.08)   | 1.57 (0.39)    | 0.17 (0.02)    | 4.33 (16.02)   | 51.14 (3.72)   | 2.72 (1.67)    |
| 5              | 82.43 (2.09)  | 105.05 (6.30) | 13.20 (1.54)  | 10.93 (0.54)  | 1.51 (0.05)    | 0.21 (0.05)    | 5.51 (16.76)   | 52.39 (3.74)   | 1.47 (1.08)    |
| 7              | 85.14 (1.96)  | 111.11 (7.44) | 13.76 (6.67)  | 14.41 (1.54)  | 2.21 (0.60)    | 0.22 (0.05)    | 6.16 (17.02)   | 52.75 (3.26)   | 0.90 (0.51)    |
| 9              | 91.34 (1.95)  | 120.06 (8.06) | 14.83 (7.16)  | 16.52 (1.76)  | 2.63 (0.71)    | 0.29 (0.07)    | 5.95 (18.09)   | 48.68 (3.23)   | 0.59 (0.40)    |
| 11             | 94.39 (1.49)  | 126.89 (8.71) | 15.23 (7.97)  | 19.99 (1.90)  | 2.72 (1.06)    | 0.34 (0.09)    | 5.37 (18.50)   | 49.12 (3.15)   | 0.40 (0.28)    |
| 13             | 94.79 (1.57)  | 130.48 (8.96) | 15.48 (8.23)  | 23.41 (2.23)  | 3.21 (1.23)    | 0.41 (0.11)    | 7.00 (18.57)   | 45.66 (2.95)   | 0.28 (0.21)    |
| 15             | 97.71 (2.07)  | 135.43 (9.99) | 16.04 (8.55)  | 25.10 (3.09)  | 3.42 (1.25)    | 0.51 (0.15)    | 7.79 (19.34)   | 41.90 (2.80)   | 0.21 (0.15)    |
| 17             | 97.39 (1.72)  | 141.98 (9.40) | 17.50 (8.47)  | 31.91 (3.14)  | 4.83 (1.51)    | 0.55 (0.12)    | 7.63 (19.38)   | 39.09 (2.72)   | 0.15 (0.11)    |
| 19             | 95.80 (2.68)  | 146.67 (10.54)| 17.57 (8.15)  | 32.55 (3.77)  | 5.26 (1.35)    | 0.69 (0.15)    | 7.34 (19.48)   | 36.64 (2.63)   | 0.11 (0.07)    |
| 21             | 98.49 (3.03)  | 147.88 (11.99)| 18.09 (8.58)  | 37.51 (6.21)  | 6.20 (1.47)    | 0.85 (0.14)    | 9.31 (20.30)   | 34.60 (2.54)   | 0.07 (0.04)    |
| 23             | 97.67 (3.49)  | 163.81 (14.45)| 21.22 (9.26)  | 53.04 (8.91)  | 8.12 (2.48)    | 1.38 (0.21)    | 9.21 (21.23)   | 37.93 (2.19)   | 0.04 (0.02)    |
| 25             | 95.24 (3.17)  | 187.68 (18.72)| 21.55 (12.90)| 81.51 (13.86)| 10.62 (4.97)   | 3.13 (0.52)    | 10.36 (21.33)  | 34.17 (2.10)   | 0.02 (0.01)    |
Table 1.11: Market Price of Risk: Contemporaneous Decomposition: This table reports the estimated coefficients from cross-sectional regressions of the liquidity-adjusted CAPM for the contemporaneous decomposition. At the end of each year, beginning in 1963 and ending in 1999, eligible stocks are sorted into 25 portfolios according to their annual Amihud (2002) illiquidity measures in Panel A and according to volatility of their daily illiquidity measures in Panel B. The breakpoints are based on eligible stocks so each portfolio has approximately the same number of stocks at formation. The individual portfolios are value-weighted and the market portfolio is equal-weighted. I report special cases of the following relationship: $E(r_i^t - r_f^t) = \alpha + \kappa E[K_i^t] + \lambda \beta + \lambda \beta^* + \lambda_{c,r} \beta_{c,r}^t + \lambda_{c,c} \beta_{c,c}^t + \lambda_{c,r} \beta_{c,r}^t + \lambda_{c,c} \beta_{c,c}^t$, where $\beta$ is the portfolio’s CAPM beta, $\beta^*$ is the portfolio’s liquidity-adjusted CAPM beta, $\beta_{c,r}$ is the portfolio’s non-liquidity news sensitivity to the market’s non-liquidity news, $\beta_{c,c}$ is the portfolio’s non-liquidity news sensitivity to the market’s liquidity news, $\beta_{c,r}$ is the portfolio’s liquidity news sensitivity to the market’s non-liquidity news, $\beta_{c,c}$ is the portfolio’s liquidity news sensitivity to the market’s liquidity news, and $\beta_c \equiv \beta_{c,c} - \beta_{c,r} - \beta_{c,c}$ is the aggregate liquidity risk beta as defined by Acharya and Pedersen (2005). The reported estimates are small-sample bias corrected using the jackknife procedure. Robust standard errors are reported in parenthesis. Both jackknife procedures take into account the pre-estimation of the news series and the betas. The $R^2$ is obtained in a single cross-sectional regression and the adjusted $R^2$ is in parenthesis.

### Panel A: Portfolios Ranked by Proportional Liquidity Levels

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$E[K]$</th>
<th>$\beta$</th>
<th>$\beta^*$</th>
<th>$\beta_{c,r}$</th>
<th>$\beta_{c,c}$</th>
<th>$\beta_c$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.618</td>
<td>1.311</td>
<td>0.744</td>
<td>0.171</td>
<td>0.171</td>
<td>0.697</td>
<td>0.569</td>
</tr>
<tr>
<td>(0.111)</td>
<td>(0.122)</td>
<td></td>
<td>(0.078)</td>
<td>(2.102)</td>
<td>(3.879)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-0.115</td>
<td>1.000</td>
<td>0.725</td>
<td>0.697</td>
<td>0.697</td>
<td>0.64</td>
<td>0.550</td>
</tr>
<tr>
<td>(0.073)</td>
<td>(0.085)</td>
<td></td>
<td>(0.257)</td>
<td>(2.102)</td>
<td>(3.879)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-0.102</td>
<td>1.057</td>
<td>0.708</td>
<td>0.835</td>
<td>0.835</td>
<td>0.820</td>
<td>0.820</td>
</tr>
<tr>
<td>(0.064)</td>
<td>(0.171)</td>
<td></td>
<td>(0.074)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-0.120</td>
<td>1.000</td>
<td>0.734</td>
<td>0.171</td>
<td>0.171</td>
<td>0.697</td>
<td>0.670</td>
</tr>
<tr>
<td>(0.068)</td>
<td>(0.078)</td>
<td></td>
<td>(2.102)</td>
<td>(3.879)</td>
<td>(3.879)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.812</td>
<td>23.906</td>
<td>-0.744</td>
<td>-259.492</td>
<td>-259.492</td>
<td>0.914</td>
<td>0.902</td>
</tr>
<tr>
<td>(0.162)</td>
<td>(3.120)</td>
<td></td>
<td>(0.257)</td>
<td>(3.879)</td>
<td>(3.879)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.476</td>
<td>1.000</td>
<td>0.215</td>
<td>293.782</td>
<td>293.782</td>
<td>0.775</td>
<td>0.731</td>
</tr>
<tr>
<td>(0.219)</td>
<td>(0.300)</td>
<td></td>
<td>(0.257)</td>
<td>(3.879)</td>
<td>(3.879)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>-0.024</td>
<td>38.605</td>
<td>0.043</td>
<td>61.554</td>
<td>61.554</td>
<td>0.916</td>
<td>0.894</td>
</tr>
<tr>
<td>(0.172)</td>
<td>(8.372)</td>
<td></td>
<td>(0.343)</td>
<td>(14.668)</td>
<td>(14.668)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Panel B: Portfolios Ranked by Liquidity Risk

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>$E[K]$</th>
<th>$\beta$</th>
<th>$\beta^*$</th>
<th>$\beta_{c,r}$</th>
<th>$\beta_{c,c}$</th>
<th>$\beta_c$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.609</td>
<td>1.285</td>
<td>0.696</td>
<td>0.596</td>
<td>0.596</td>
<td>0.581</td>
<td>0.598</td>
</tr>
<tr>
<td>(0.122)</td>
<td>(0.136)</td>
<td></td>
<td>(0.075)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-0.097</td>
<td>1.000</td>
<td>0.696</td>
<td>0.596</td>
<td>0.596</td>
<td>0.579</td>
<td>0.596</td>
</tr>
<tr>
<td>(0.066)</td>
<td>(0.075)</td>
<td></td>
<td>(0.075)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-0.024</td>
<td>1.397</td>
<td>0.591</td>
<td>0.778</td>
<td>0.778</td>
<td>0.758</td>
<td>0.758</td>
</tr>
<tr>
<td>(0.057)</td>
<td>(0.184)</td>
<td></td>
<td>(0.068)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-0.035</td>
<td>1.000</td>
<td>0.608</td>
<td>5.429</td>
<td>5.429</td>
<td>0.617</td>
<td>0.582</td>
</tr>
<tr>
<td>(0.062)</td>
<td>(0.074)</td>
<td></td>
<td>(3.014)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-0.146</td>
<td>-1.579</td>
<td>0.786</td>
<td>32.388</td>
<td>32.388</td>
<td>0.778</td>
<td>0.778</td>
</tr>
<tr>
<td>(0.091)</td>
<td>(2.465)</td>
<td></td>
<td>(2.903)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.118</td>
<td>1.000</td>
<td>0.469</td>
<td>62.425</td>
<td>62.425</td>
<td>0.640</td>
<td>0.640</td>
</tr>
<tr>
<td>(0.203)</td>
<td>(0.403)</td>
<td></td>
<td>(1.343)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1.555</td>
<td>-54.569</td>
<td>-0.097</td>
<td>564.639</td>
<td>564.639</td>
<td>0.834</td>
<td>0.834</td>
</tr>
<tr>
<td>(0.448)</td>
<td>(9.411)</td>
<td></td>
<td>(5.699)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 1.12: Proportional Cost Decomposition: This table reports the estimated coefficients from cross-sectional regressions of the liquidity-adjusted CAPM for the proportional cost decomposition. At the end of each year, beginning in 1963 and ending in 1999, eligible stocks are sorted into 25 portfolios according to their annual Amihud (2002) illiquidity measures in Panel A and according to volatility of their daily illiquidity measures in Panel B. The breakpoints are based on eligible stocks so each portfolio has approximately the same number of stocks at formation. The individual portfolios are value-weighted and the market portfolio is equally-weighted. I report special cases of the following relationship:

\[ E(r_i^t - r_f^t) = \alpha + \kappa E[K]_t^i + \lambda \beta + \lambda E \beta^* + \lambda \beta_{r,c} + \lambda \beta_{c,r} + \lambda \beta_{c,c} + \lambda \beta_{c,r} \beta_{c,c} + \lambda \beta_{c,c} \beta_{c,r} \]

where \( \beta \) is the portfolio’s liquidity beta, \( \beta^* \) is the portfolio’s liquidity-adjusted CAPM beta, \( \beta_{r,c} \) is the portfolio’s non-liquidity news sensitivity to the market’s non-liquidity news, \( \beta_{c,r} \) is the portfolio’s liquidity-adjusted CAPM beta, \( \beta_{c,c} \) is the aggregate liquidity risk beta as defined by Acharya and Pedersen (2005). The reported estimates are small-sample bias corrected using the jackknife procedure. Robust standard errors are reported in parenthesis. Both jackknife procedures take into account the pre-estimation of the news series and the betas. The \( R^2 \) is obtained in a single cross-sectional regression and the adjusted \( R^2 \) is in parenthesis.

### Panel A: Portfolios Ranked by Proportional Liquidity Levels

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>( E[K] )</th>
<th>( \beta )</th>
<th>( \beta^* )</th>
<th>( \beta_{r,c} )</th>
<th>( \beta_{c,r} )</th>
<th>( \beta_{c,c} )</th>
<th>( \beta_e )</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.618</td>
<td>1.311</td>
<td>0.940</td>
<td>-0.205</td>
<td>-36.988</td>
<td>2.162</td>
<td>0.064</td>
<td>0.569</td>
</tr>
<tr>
<td>2</td>
<td>-0.391</td>
<td>1.00</td>
<td>0.940</td>
<td>-0.205</td>
<td>-36.988</td>
<td>2.162</td>
<td>0.064</td>
<td>0.550</td>
</tr>
<tr>
<td>3</td>
<td>-0.297</td>
<td>1.394</td>
<td>0.877</td>
<td>-0.205</td>
<td>-36.988</td>
<td>2.162</td>
<td>0.064</td>
<td>0.500</td>
</tr>
<tr>
<td>4</td>
<td>-0.340</td>
<td>1.00</td>
<td>0.940</td>
<td>-0.205</td>
<td>-36.988</td>
<td>2.162</td>
<td>0.064</td>
<td>0.500</td>
</tr>
<tr>
<td>5</td>
<td>0.420</td>
<td>10.559</td>
<td>-0.205</td>
<td>-36.988</td>
<td>2.162</td>
<td>0.064</td>
<td>0.500</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.150</td>
<td>10.00</td>
<td>0.551</td>
<td>225.231</td>
<td>12.169</td>
<td>956.514</td>
<td>0.713</td>
<td>0.827</td>
</tr>
<tr>
<td>7</td>
<td>0.367</td>
<td>13.040</td>
<td>0.155</td>
<td>382.647</td>
<td>41.951</td>
<td>-21e+03</td>
<td>0.879</td>
<td>0.847</td>
</tr>
</tbody>
</table>

### Panel B: Portfolios Ranked by Liquidity Risk

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>( E[K] )</th>
<th>( \beta )</th>
<th>( \beta^* )</th>
<th>( \beta_{r,c} )</th>
<th>( \beta_{c,r} )</th>
<th>( \beta_{c,c} )</th>
<th>( \beta_e )</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.699</td>
<td>1.285</td>
<td>0.806</td>
<td>0.936</td>
<td>4.450</td>
<td>0.755</td>
<td>0.577</td>
<td>0.598</td>
</tr>
<tr>
<td>2</td>
<td>-0.325</td>
<td>1.00</td>
<td>0.920</td>
<td>0.806</td>
<td>0.936</td>
<td>4.450</td>
<td>0.755</td>
<td>0.581</td>
</tr>
<tr>
<td>3</td>
<td>-0.179</td>
<td>1.516</td>
<td>0.731</td>
<td>0.806</td>
<td>0.936</td>
<td>4.450</td>
<td>0.755</td>
<td>0.533</td>
</tr>
<tr>
<td>4</td>
<td>-0.232</td>
<td>1.00</td>
<td>0.806</td>
<td>0.936</td>
<td>4.450</td>
<td>0.755</td>
<td>0.577</td>
<td>0.749</td>
</tr>
<tr>
<td>5</td>
<td>-0.315</td>
<td>0.314</td>
<td>0.936</td>
<td>0.806</td>
<td>0.936</td>
<td>4.450</td>
<td>0.755</td>
<td>0.726</td>
</tr>
<tr>
<td>6</td>
<td>-0.880</td>
<td>1.00</td>
<td>0.886</td>
<td>116.634</td>
<td>-11.586</td>
<td>-2.2e+03</td>
<td>0.574</td>
<td>0.720</td>
</tr>
<tr>
<td>7</td>
<td>-1.824</td>
<td>0.915</td>
<td>2.906</td>
<td>23.336</td>
<td>-15.802</td>
<td>-3.6e+03</td>
<td>0.770</td>
<td>0.709</td>
</tr>
</tbody>
</table>
Table 1.13: Fixed Cost Decomposition: This table reports the estimated coefficients from cross-sectional regressions of the liquidity-adjusted CAPM for the fixed cost decomposition. At the end of each year, beginning in 1963 and ending in 1999, eligible stocks are sorted into 25 portfolios according to their annual Amihud (2002) illiquidity measures in Panel A and according to volatility of their daily illiquidity measures in Panel B. The breakpoints are based on eligible stocks so each portfolio has approximately the same number of stocks at formation. The individual portfolios are value-weighted and the market portfolio is equal-weighted. I report special cases of the following relationship:

\[ E(r_i - r_f) = \alpha + \beta \beta^* + \lambda \beta_{cr} + \lambda_{rc} \beta_{r,c} + \lambda_{r,c} \beta_{r,c} + \lambda_{r,r} \beta_{r,r} + \lambda_{c,c} \beta_{c,c} + \lambda \beta_{c} \]

where \( \beta \) is the portfolio’s CAPM beta, \( \beta^* \) is the portfolio’s liquidity-adjusted CAPM beta, \( \beta_{cr}, \beta_{r,c}, \beta_{c,c} \) are the portfolio’s non-liquidity news sensitivity to the market’s non-liquidity news, \( \beta_{r,c} \) is the portfolio’s non-liquidity news sensitivity to the market’s liquidity news, \( \beta_{c,c} \) is the portfolio’s liquidity news sensitivity to the market’s non-liquidity news, and \( \beta_{c} \) is the aggregate liquidity risk beta as defined by Acharya and Pedersen (2005). The reported estimates are small-sample bias corrected using the jackknife procedure. Robust standard errors are reported in parenthesis. Both jackknife procedures take into account the pre-estimation of the news series and the betas. The \( R^2 \) is obtained in a single cross-sectional regression and the adjusted \( R^2 \) is in parenthesis.

### Panel A: Portfolios Ranked by Proportional Liquidity Levels

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>( E[K] )</th>
<th>( \beta )</th>
<th>( \beta^* )</th>
<th>( \beta_{cr} )</th>
<th>( \beta_{r,c} )</th>
<th>( \beta_{c,c} )</th>
<th>( \beta_{c} )</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>−0.618</td>
<td>1.311</td>
<td>0.569</td>
<td>(0.111)</td>
<td>(0.122)</td>
<td>0.550</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>−0.391</td>
<td>1.000</td>
<td>1.000</td>
<td>0.617</td>
<td>(0.090)</td>
<td>(0.100)</td>
<td>0.600</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>−0.297</td>
<td>1.394</td>
<td>0.877</td>
<td>0.811</td>
<td>(0.072)</td>
<td>(0.150)</td>
<td>0.794</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>−0.214</td>
<td>1.000</td>
<td>0.732</td>
<td>0.561</td>
<td>(0.074)</td>
<td>(0.098)</td>
<td>0.678</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.056</td>
<td>−1.428</td>
<td>0.218</td>
<td>−0.775</td>
<td>(0.190)</td>
<td>(0.352)</td>
<td>0.838</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.075</td>
<td>1.000</td>
<td>0.725</td>
<td>14.446</td>
<td>(0.156)</td>
<td>(0.425)</td>
<td>0.701</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>−0.074</td>
<td>2.075</td>
<td>0.920</td>
<td>14.273</td>
<td>(1.687)</td>
<td>(1.619)</td>
<td>0.865</td>
<td></td>
</tr>
</tbody>
</table>

### Panel B: Portfolios Ranked by Liquidity Risk

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>( E[K] )</th>
<th>( \beta )</th>
<th>( \beta^* )</th>
<th>( \beta_{cr} )</th>
<th>( \beta_{r,c} )</th>
<th>( \beta_{c,c} )</th>
<th>( \beta_{c} )</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>−0.609</td>
<td>1.285</td>
<td>0.598</td>
<td>(0.122)</td>
<td>(0.136)</td>
<td>0.581</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>−0.325</td>
<td>1.000</td>
<td>0.920</td>
<td>0.533</td>
<td>(0.086)</td>
<td>(0.097)</td>
<td>0.513</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>−0.179</td>
<td>1.517</td>
<td>0.730</td>
<td>0.749</td>
<td>(0.064)</td>
<td>(0.173)</td>
<td>0.726</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>−0.097</td>
<td>1.000</td>
<td>0.584</td>
<td>0.394</td>
<td>(0.077)</td>
<td>(0.106)</td>
<td>0.580</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>−0.018</td>
<td>−0.270</td>
<td>0.408</td>
<td>0.394</td>
<td>(0.154)</td>
<td>(1.817)</td>
<td>0.542</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>−0.019</td>
<td>1.000</td>
<td>1.078</td>
<td>5.931</td>
<td>(0.179)</td>
<td>(0.531)</td>
<td>0.726</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.070</td>
<td>−4.553</td>
<td>0.695</td>
<td>−8.827</td>
<td>(0.219)</td>
<td>(3.858)</td>
<td>0.775</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 2

Surprise! It’s Illiquid: The January 2006 Tokyo Stock Exchange Shutdown as a Natural Experiment

2.1 Introduction

On January 18, 2006, the Tokyo Stock Exchange (TSE) unexpectedly closed twenty minutes early. Forty minutes earlier, investors were informed that the number of transactions on that day, approximately 4.5 million, had reached the exchange’s transaction processing system capacity. Further, the exchange disclosed that upgrading its technology could take six to twelve months, that the exchange would shut down automatically on any day in which the daily capacity had been reached, and that the exchange would close half an hour early each day in the near future to reduce the risk of another unanticipated shutdown. The event provides a unique opportunity to learn about the value investors place on liquidity. The market shutdown was not simply a single passing disruption to the system. The unexpected market closure, coupled with the information disclosed by the exchange during and after the event, likely led investors to update their beliefs about their portfolios’ liquidity risk. Stocks are not necessarily less liquid – it is entirely possible that the threshold will not be reached before the systems are updated. However, the estimated probability of such an event occurring has increased. Hence, the liquidity risk of holding an asset traded on
the TSE has increased. In addition, the higher probability of an illiquidity event results in higher expected future illiquidity for all stocks listed on the exchange. Second, the perceived change in market liquidity is systematic to all TSE traded assets. Where trading volume originates is irrelevant; once a certain threshold is met, trading ceases for all assets listed on the exchange. Finally, the increased liquidity risk is exogenous in the sense that investors may be fully ready and willing but simply unable to provide liquidity.

Using transactions data and the Gibbs sampler estimator suggested by Hasbrouck (2006), I estimate the bid-ask spread for all stocks traded on the TSE over three ten trading day periods: the ten days prior to the closure, the ten day period including and after the closure, and the ten day period after the second window. The average spread across all stocks listed on the TSE increased from 23.6 basis points during the first period to 26.2 basis points during the period after the closure, a statistically significant result. The increase in spreads was slightly larger (3.1 basis points) for stocks that are only traded on the TSE. Estimated spreads during the third window provide evidence that spreads reverted to their levels prior to the market closure; the difference in the mean spread between the first and third period is not statistically significant. The findings suggest that spreads only partially capture the change in liquidity associated with the increased risk of market closure because the increased liquidity risk had not subsided by the third period.

Was the TSE likely to experience additional market closures after the January 18 event? After estimating an AR(1) process for the number of equity transactions processed at the daily frequency using transaction level data for the TSE with sample period beginning on October 1, 2005, I perform a Monte Carlo analysis and simulate the number of daily transactions over a six month horizon beginning on January 19, 2006. I estimate the probability of at least one addition market closure during the six month period to be approximately 70 percent and the expected number of closures to be 1.5. The simulations, which are performed using the information available on January 18, 2006, indicate that investors had reason to suspect that their future ability to trade was lower than previously anticipated.

Because the market closure on January 18 led investors to update their beliefs about the liquidity risks they face, if investors require a premium for holding assets with liquidity
risk, then asset prices should change to reflect the new information. I investigate the cross-sectional variation in asset price changes subsequent to the event. The first component of the analysis tests for a “flight-to-quality” by comparing the returns of the largest capitalization quintile of stocks on the TSE to the smallest quintile after controlling for industry effects. Large stocks immediately experienced a statistically significant 3 percent capital gain relative to small stocks. The shift in value is persistent and increases to 4 percent over a ten day period. Because two events, the announced Livedoor investigation and the market closure, occurred on January 18, it is not clear how much of the flight-to-quality is due to updated beliefs about future liquidity versus concern about fraud.

In the second component of my analysis, I compare the post-event returns of 657 stocks that are traded only on the TSE to the returns of 297 stocks that are traded on the TSE and at least one other Japanese exchange. The liquidity of stocks in the latter group are less affected by a TSE market closure because they have an addition venue for trade. I document an immediate statistically significant relative loss in value of 0.5 percent for the TSE only category. The relative capital loss persists and increases to approximately 2.0 percent over a ten day period. Next, I compare cross-listed stocks with large amounts of trade conducted on OSE or NSE against those with the majority of their volume traded on the TSE. I find no statistically significant shifts in relative value across the two groups. Together, the findings suggest that it is the ability to trade elsewhere and not how trade is split amongst exchanges that was important to investors.

The final component of my analysis investigates the relative value change for stocks with different liquidity levels. After sorting stocks on a variety of liquidity-related characteristics, such as turnover, monetary volume, and the Amihud (2002) illiquidity measure, and controlling for size and industry effects, I form liquidity portfolios that are long the most liquid assets and short the least liquid assets. Sorting on different liquidity characteristics provides similar results; liquid stocks lost value relative to illiquid stocks. For instance, sorting on turnover, the liquidity portfolio has an immediate capital loss of 5 percent that increases to 6 percent over a ten day period. Sorting on the Amihud (2002) illiquidity measure, the liquidity portfolio loses approximately 2.5 percent of its value on January 18. The capital
loss increases to almost 4 percent by the end of the month.

My analysis is related to two strings in the literature on the relationship between returns and time-series variation in liquidity. The first set considers value changes due to liquidity events. Amihud, Mendelson, and Lauterbach (1997) investigate the exogenous transfer of stocks listed on the Tel Aviv Stock Exchange from a once-a-day call auction to semi-continuous trading. He finds that transferred stocks realized a persistent capital gain of 5-6 percent. In addition, stocks that experienced greater improvements in their liquidity had larger capital gains. Muscarella and Piwowar (2001) consider a similar transfer of stocks from call auctions to a continuous market on the Paris Bourse and reports that the improvements in liquidity were associated with higher valuations. Hegde and McDermott (2003) investigate the effect of exogenous changes in asset liquidity due to joining or leaving the S&P 500 Index. They document that liquidity, as measured by Kyle’s λ and the bid-ask spread, increases upon inclusion and that the capital gain that occurs subsequent to joining the index is positively related to the improved liquidity. The three papers are similar because they investigate returns associated with idiosyncratic liquidity events. My paper is different for two reasons. First, the event that I study is a systematic change in liquidity. Second, I do not investigate how the valuation of the entire market changed because of the event. I study how relative valuations for assets with differing liquidity characteristics were impacted by the liquidity event.

The second set of literature investigates the relationship between time-series variation in aggregate liquidity and expected returns. Rather than studying a particular liquidity event, the papers in this area construct a time series for a measure of aggregate liquidity and consider the covariation between returns and aggregate liquidity. Amihud (2002) finds that stocks’ excess returns increase in expected illiquidity and decrease in unexpected illiquidity. He finds that smaller firms have returns that are more sensitive to changes in aggregate liquidity. Pastor and Stambaugh (2003) propose and document that stocks expected returns increase in their return sensitivities to changes in market liquidity. Acharya and Pedersen (2005) propose a Liquidity-Adjusted CAPM that relates expected excess returns to expected liquidity levels, systematic gross return risk, and systematic liquidity risk. One of the
three sources of systematic liquidity risk in their model is the asset’s return sensitivity to changes in aggregate liquidity. They report that less liquid stocks have higher exposure to liquidity risk. My paper is similar to these papers because I investigate cross-sectional variation in asset return sensitivities to changes in market liquidity. Unlike the papers in this area, I do not form an aggregate measure of liquidity and I do not relate the cross-sectional differences to expected returns. I study the cross-sectional variation in asset return sensitivities to the single liquidity event that transpired on January 18. Interestingly, my findings are not consistent with those reported by Amihud (2002) and Acharya and Pedersen (2005). Both papers document that illiquid stocks have returns with greater sensitivities to changes in aggregate liquidity, which is consistent with a “flight-to-liquidity”, the idea that when liquidity dries up, liquid stocks outperform illiquid stocks. For the January 18 event studied in this paper, I find that illiquid stocks, as inferred by turnover, monetary volume, and number of transactions or as identified by the Amihud (2002) price-impact illiquidity measure, experience a statistically significant and persistent capital loss of 3 to 8 percent after adjusting for size and industry effects. This finding suggests that not all negative shocks to aggregate liquidity lead to a flight-to-liquidity.

The rest of the paper is organized as follows: The following section presents a brief history of the event examined in the paper. Section 3 describes the data, the econometric specification, and reports the results. Section 4 concludes.

2.2 Background Information

2.2.1 The Exchanges

Japan has five stock exchanges: Tokyo Stock Exchange (TSE), Osaka Securities Exchange (OSE), Nagoya Stock Exchange (NSE), Fukuoka Stock Exchange, and Sapporo Securities Exchange. The first three are the largest in terms of market capitalization and monetary trading volume, and the TSE, the second largest stock exchange in the world by monetary volume, dominates in both categories. Almost all the research of the Japanese stock market
focuses exclusively on the TSE and the summary statistics described below provide evidence on why this is the case.

Stocks may be traded on more than one exchange. Table 2.1 presents summary statistics for the seven permutations of exchange listings using the TSE, OSE, and NSE. The first three rows are for stocks listed on only one exchange, the next three are for stocks listed on two exchanges, and the final row is for stocks traded on the three exchanges. The next section provides a complete description of the data. My sample includes a total of 2847 stocks, 2370 of which are listed on the TSE. The majority of stocks listed on the TSE are not listed on either of the other two exchanges, but the majority of stocks listed on the NSE or OSE are also listed on the TSE. For instance, 83 percent of stocks traded on the OSE are also traded on TSE, but only 29 percent of stocks traded on the TSE are also traded on the OSE. Stocks traded on the TSE have a mean market capitalization of ¥216 billion compared to ¥19 billion for those not traded on the TSE. The market capitalization of stocks listed on both the TSE and OSE is approximately twice that of those listed only on the TSE. Stocks listed on all three exchanges are about 12 times as large as those listed only on the TSE. The 83 percent of stocks traded on the TSE represent 98 percent of the Japanese stock market capitalization.

Although the OSE and NSE appear to be relatively unimportant when compared against the TSE and are generally overlooked by most research, I include the two exchanges in my analysis. Aside from the cross-sectional variation in liquidity amongst the stocks, 1552 stocks are only tradable on the TSE, while 818 stocks are cross-listed on at least one additional Japanese exchange. A TSE market closure disproportionately affects the liquidity of stocks in the two categories and the first step of my analysis is to compare the returns of the stocks in the two groups subsequent to the January 18 liquidity event. An alternate approach is to compare stocks traded on the TSE to those that are not. I do not perform this analysis for two reasons. First, the total market share of stocks not traded on the TSE is almost insignificant at less than 2 percent. Second, the selection bias, as evidenced by the differences in mean market capitalization, is likely to be substantial.
2.2.2 The Liquidity Event

In January 2006, the TSE was capable of executing close to 4.5 million trades per day, approximately 50 percent more than the average number of daily trades in December 2005. Late Monday night on January 16, prosecutors raided the offices of Livedoor, an upstart Internet portal that had been popular with small investors. Two days later, the Japanese media reported that Livedoor was suspected of hiding a ¥1 billion loss, or $8.7 million, by claiming revenue from three affiliated companies. Lacking buy orders, Livedoor shares did not trade Wednesday and were marked down by the daily ¥100 limit to ¥496. Many stocks listed on the TSE began to experience a surge in the number of sell orders. By lunchtime, the president of the TSE announced that the exchange processed 2.3 million transactions in morning trading and that the exchange’s capacity would likely be hit. At 2:00 pm, an hour before the market’s usual closing time, normally “instantaneous” trades took over 5 minutes to process and the exchange formally announced that trading would cease 20 minutes early.

After shutting down the exchange early, the TSE announced that the daily trading period would be shortened by a half-hour until the trading system was upgraded. Investors were given mixed signals; they were warned that exceeding the 4.5 million trade capacity before the upgrades were complete would lead to another shutdown and reassured that another shutdown due to excessive volume was almost impossible. On February 24, 2006 the TSE announced its plans to resume normal trading hours no later than May. The expected improvements would allow the TSE to process seven million executions per day.

The event is interesting for a number of reasons. First, the liquidity shock is systematic. Where and why TSE transactions originate is unimportant. Once a certain threshold is reached, all trading on the TSE ceases. For the 1552 stocks listed only on the TSE, a market closure equates to almost perfect illiquidity. The remaining 818 stocks listed on the TSE are also listed on another Japanese exchange and their liquidity levels are not as vulnerable to a shutdown. Second, a future market shutdown is independent of market direction. Unlike circuit breakers set by exchanges that limit the daily downward price movement of a stock or index, in this case, price direction does not matter. Any news, good
or bad, that increases trading activity also increases the probability of market shutdown. Finally, the liquidity event is exogenous. Liquidity does not dry up because of increased uncertainty or maximum allowable downward price movement. Liquidity providers may be fully ready and willing, but simply unable to trade. Indeed, a large number of noise traders, which are generally credited for increased market liquidity, may be the cause of market shutdown.

2.3 Empirical Analysis

2.3.1 Data

I analyze stocks listed on the TSE, NSE, and OSE. The list of listed companies for each exchange is provided by the exchanges’ respective websites. Transaction level data over the period October 1, 2005 through February 28, 2006 is obtained from the TSE for all of its traded stocks. Each observation includes the price, time, and trade size. The number of shares outstanding is a snapshot taken on January 31, 2006 from the Marketwatch.com website maintained by Dow Jones. Thus, summary statistics for market capitalization provided in Table 2.1 are not adjusted for stock splits are any other time-variation in the number of outstanding shares. Nikkei 225 Index levels, which I use to proxy for the Japanese aggregate portfolio, and price and volume information used to calculate summary statistics in Table 2.1 for the 477 stocks not traded on the TSE are also obtained from the Marketwatch.com. Industry classification for each stock listed on the TSE is obtained from the exchange’s website and is a snapshot taken on November 10, 2006. The industry classification of the 54 stocks that were traded on the TSE at the end of 2005, but were no longer listed on November 10, 2006 is obtained from Bloomberg.com.

2.3.2 Likelihood of Future TSE Shutdown

The analysis in this paper presupposes that investors revised their expectations of the likelihood of future market shutdown. Because prices are a forward looking measure of
value, if the event is not informative about future ability to trade, there is little reason to suspect that any asymmetric changes in value follow the event. In this section, I estimate the likelihood of future shutdown.

Figure 2.1 plots the total number of equity transactions on each day for my sample period. The spike in transactions that led to the shutdown is clearly visible on January 18, 2006. I calculate the number of equity transactions on that day to be 2.03 million. However, it is the total number of transactions cleared by the exchange, including those in the bond and derivative markets, that is constrained by TSE's computer systems. Unfortunately, my data does not include non-equity transactions. To predict the probability that the TSE capacity has been reached, I assume that the proportion of equity to total transactions is constant over the horizon under investigation. With the capacity of the exchange on January 18 at approximately 4.5 million, on that day equity transactions represented approximately 45 percent of total TSE transactions. The Japanese press reported the average number of transactions in December 2005 to be approximately 3 million per day. I estimate the average number of equity transactions during the same month to be 1.46 million per day or approximately 49 percent of the total. The similarity between the two estimates provides some support for my constant proportion assumption.

I further assume the number of transactions, in millions and denoted by $\tau_t$, follow an AR(1) process and estimate the model via OLS with the sample up to and including the event date:

$$
\tau_{t+1} = 0.5065 + 0.6363\tau_t + \epsilon_{t+1}, \\
R^2 = 0.3582 \quad \sigma(\epsilon) = 0.1879.
$$

$$
(0.1041) (0.1442)
$$

Standard errors are in parenthesis. Including a second lag provides a negligible improvement in fit and the AR(2) coefficient is not statistically significant. Using equation (2.1), I simulate 25 million 120 day (approximately six months) paths for the number of daily transactions. The shocks used in the simulations are the residuals from regression (2.1) resampled with replacement. I include both signs of each residual to double the error sample size from 70 to 140. I resample the residuals to avoid making any distributional
assumptions on the error process that may remove the ‘outliers’ that are likely responsible for the spike in transactions. Predicted market closure occurs when the simulated number of transactions exceeds 2.03 million, the number of equity transactions on January 18, 2006.

Figure 2.2 plots the histogram for the simulations. The expected number of shutdowns over the six month period following the January 18 shutdown is approximately 1.5. The estimated probabilities of zero and one shutdown, respectively, are 31 and 29 percent. There is a 21 percent chance of three or more shutdowns, at least one shutdown every two months. The analysis suggests that investors have reason to believe that future shutdowns are not only possible, but also probable. Hence, any negative value associated with the new information about future potential liquidity events may be reflected in price movements subsequent to the January 18 market closure.

2.3.3 Relative Value Changes

This section investigates the effect of the January 18 shutdown on relative valuations. The general strategy for the analysis is summarized by the following outline:

1. For each asset, I estimate the cumulative simple return obtained by purchasing the stock on January 17, 2006 and selling the stock on each of the remaining 10 trading days in January.

2. Assets are then sorted on liquidity-related characteristics, such as turnover and Amihud (2002)’s illiquidity measure.

3. Zero-cost portfolios are formed that are long and short, respectively, the liquid and illiquid quintile from the previous step sort.

4. Finally, I estimate the relative value change of the liquidity portfolios over the 10 day period.

For the first step, I estimate the cumulative simple return with and without adjusting for a market factor, which I represent by the Nikkei 225 Index. The former approach is denoted by “market-adjusted cumulative returns” and the latter is “unadjusted.” There is
evidence that factor models may not be appropriate for explaining Japanese stock returns. Research by Hawawini (1991) indicates that the market factor cannot explain Japanese cross-sectional stock returns. Daniel, Titman, and Wei (2001) show that Japanese stock returns are more closely related to their book-to-market ratios than United States stocks and reject the Fama and French (1993) three-factor model. Hu (1997), who investigates the impact of liquidity as measured by asset turnover on asset pricing on the TSE, does not include the market factor in regressions. I include the market-adjusted cumulative returns for robustness.

The cumulative simple return for each stock is estimated via the following OLS regression,

\[ r_{it}^i = \alpha_i + \beta_i r_{mt}^m + \sum_{d=1}^{10} \alpha_d^i \psi_{d,t} + \epsilon_{it}, \]  

(2.2)

where \( r_{it}^i \) is the daily log return for asset \( i \) at time \( t \), \( r_{mt}^m \) is the log return for the Nikkei 225 Index, \( \psi_{d,t} \) is 1 on the \( d \)th trading day after January 17, 2006 and 0 otherwise. The cumulative simple return is then

\[ R_{id}^i \equiv \exp \left( \sum_{d=1}^{10} \alpha_d^i \right) - 1. \]  

(2.3)

Further detail on the second and third steps are provided in the relevant sections. The output of the two steps are a set of \( N_L \) and \( N_S \) assets, respectively, that are the long and short components of the portfolio of interest and their market-adjusted and unadjusted cumulative returns. The cumulative return for the portfolio is calculated to be:

\[ R_{id}^p = \frac{1}{N_L} \sum_{i=1}^{N_L} R_{id,L}^i - \frac{1}{N_S} \sum_{i=1}^{N_S} R_{id,S}^i, \]  

(2.4)

where \( R_{id,L}^i \) is the return of asset \( i \) within the long component of the portfolio and \( R_{id,S}^i \) is similarly defined for the short component. I provide two metrics for testing the statistical significance of the portfolio’s cumulative returns. The first metric is the standard error of the portfolio’s cumulative return on each day and is used to test whether the relative change in value of the portfolio is significantly different than zero. The standard error is calculated
to be
\[ se\left( R_d^{\text{pop}} \right) = \left[ N_L^{-\frac{3}{2}} \sum_{i=1}^{N_L} (R_{d,L}^i - \bar{R}_{d,L})^2 + N_S^{-\frac{3}{2}} \sum_{i=1}^{N_S} (R_{d,S}^i - \bar{R}_{d,S})^2 \right]^{\frac{1}{3}}, \] (2.5)

where \( \bar{R}_{d,L} \) and \( \bar{R}_{d,S} \) are the equal-weight cumulative returns, respectively, for the \( N_L \) and \( N_S \) assets that are purchased and sold. The second metric is the standard error of a randomly-formed portfolio with the same structure – long and short \( N_L \) and \( N_S \) randomly selected stocks. The metric allows for testing whether the liquidity portfolio’s return is statistically different than a similarly structured randomly-formed portfolio. The standard error is calculated to be
\[ se\left( R_d^{\text{rand}} \right) = \left[ \frac{N_L + N_S}{N_L N_S} \cdot se\left( R_d^{\text{pop}} \right) \right], \] (2.6)

where \( se\left( R_d^{\text{pop}} \right) \) is the standard error of the \( d \)-day cumulative return for the population being sampled from when forming the benchmark portfolio. The derivation of equation (2.6) is included in the Appendix.

Industry Effects

The market closure was not the only event on January 18 that may have lead to asymmetric changes in valuation. Indeed, the closure was because of a surge in transactions due to new information about alleged fraudulent activity by one of Japan’s most popular internet companies. It is likely that the cause of the spike in trading activity may itself lead to shifts in market valuations across industries.

In addition to exchange-traded funds (ETF) and real estate investment trusts (REIT), TSE categorizes stocks according to 33 industry classifications. Table 2.2 details the number of stocks within each industry classification for the 1552 stocks that are only traded on the TSE. I aggregate industry classifications into 11 groups. For instance, the first group contains the five industries related to transportation and the second group includes the five finance related industries. As will soon be clarified, I sort stocks according to certain asset characteristics within each industry group. The reason for the aggregation is to ensure
that each group contains a sufficient number of stocks, which is not the case for individual industries.

Table 2.3 reports the cumulative returns estimated using the unadjusted model for the 11 industry groups. Every industry group suffered a capital loss on January 18. The information/communication group, which is the classification for Livedoor, had the largest capital loss, 7.9 percent, followed by services with a 7.8 percent loss. By the end of the month, as shown by $R_{10}$, every group except for wholesale/retail, services, and information/communication had fully recovered the initial capital loss on January 18. The machinery group went from an initial capital loss of 6.5 percent to a 3.25 percent capital gain.

These results suggest that over the period of interest, there were industry effects that may influence the analysis. If, for example, stocks within the information/communication group have different liquidity characteristics than those in construction, then sorting on liquidity-related attributes and examining the return of the liquidity portfolio may falsely attribute shifts in industry valuation to those related to liquidity.

For the remainder of the paper, when forming zero-cost portfolios, I control for industry in the following manner. I sort assets within each industry group on the characteristic of interest. A portfolio is then formed by buying and selling an approximately equal number of stocks from each industry. For instance, if portfolios are formed subsequent to a quintile ranking, then approximately 30 stocks within the services group will be in the long and short portfolios. The total proportion of stocks from each industry contained in the final industry-neutral portfolio reflects their relative contribution to the market portfolio, and the industry effect within the long portfolio is offset by an equal representation within the short portfolio.

**Size Effects**

Both events on January 18, the announced Livedoor investigation and the market closure, may lead to a flight-to-quality. If quality and size are positively related, then the flight-to-
quality may be tested by looking for a size effect. I form industry-adjusted quintile-ranked portfolios after sorting TSE only stocks on their respective time-series average daily market capitalization over the period October 1, 2005 to December 31, 2005. The final portfolio is long the 309 firms with the highest market capitalization and short the 307 stocks with the lowest market capitalization.

Table 2.5 reports the cumulative returns for the market adjusted and unadjusted models and Figure 2.5 plots the returns. There is an immediate and persistent shift in relative value towards larger firms. In the market adjusted model, large-cap stocks have an initial 4 percent capital gain relative to small-cap stocks. The initial capital gain drops to 3 percent over the ten day holding period. The unadjusted results are similar with an initial relative capital gain of 3 percent that grows to 4 percent over the ten day period.

Although the results suggest a flight-to-quality through size, it is not clear in what proportion the two events are responsible for the shift in value. Because of the evident size effects during the period of interest, I control for size in addition to industry. I sort all assets listed on the TSE within each industry group by their respective average daily market capitalization over the period October 1, 2005 through December 31, 2005 and remove the smallest and largest 30 percent. Thus, each industry’s contribution to the final portfolio is unaffected by the size control and only the 40 percent of stocks that are mid-cap are included in the analysis. The final industry-neutral portfolio should be less impacted by the size effects described above.

**Cross-Listed Stocks**

Of the 2370 stocks listed on the TSE, 818 are also traded on at least one of the other Japanese exchanges. The ability to trade these stocks are less influenced by an unanticipated TSE closure than those that are only traded on the TSE. Did stocks that are only tradable on the TSE lose value with respect to those with another avenue for trade? I answer the question by forming a zero-cost portfolio that is long assets traded only on the TSE and short those that are cross-listed. The portfolio is controlled for size by only including mid-cap stocks as
detailed in the previous section. The size-adjusted portfolio is long and short, respectively, 657 and 297 stocks.

Table 2.6 reports and Figure 2.4 plots the cumulative returns for the portfolio. Because the long and short components have almost exactly the same market factor loading, the results of the market adjusted and unadjusted models are nearly identical. TSE only stocks lost approximately 0.5 percent of their value relative to their cross-listed counterparts on the day of the event and the capital loss increased to approximately 2 percent over the next ten days. The capital loss is statistically significant on each day, except for the sixth day, and the portfolio’s return is significantly different than that of a portfolio that is long and short, respectively, 657 and 297 randomly selected TSE listed stocks. The results suggest that investors place a slightly higher premium on the ability to trade assets on another exchange after the January 18 market closure. Relating these relative value shifts to the 1.5 expected number of market closures estimated earlier in the paper allows for the calculation of a “premium” per expected closure; the premium for the ability to trade on an additional exchange was approximately 33 basis points per expected closure on the day of the event and increased to 133 basis points per expected closure by the end of the ten-day window.

In addition to considering the binary classification of single versus cross-listed, I also consider portfolios formed by sorting cross-listed stocks by the share of volume traded outside of the TSE. Volume data for stocks traded on OSE and NSE during the months of October through December 2005 are collected from the respective exchanges’ websites. The volume of trade conducted on the TSE during the fourth quarter of 2005 is computed using transactions data provided by the TSE. I estimate the percentage of total trading volume transacted outside of the TSE for the 818 cross-listed stocks. Table 2.4 reports the cumulative density function for the share of trading volume cleared outside of TSE. Interestingly, 7.6 percent of cross-listed stocks had no trade on the OSE or NSE during the three-month period. Approximately 80 percent of cross-listed stocks had only 1 percent of total volume clear on either OSE or NSE. Only 58 stocks or 7 percent of the cross-listed stocks had less than half of their trading volume occur on the TSE.

Were the valuations of stocks primarily traded on the OSE or NSE less influenced by the
market closure than those primarily traded on the TSE? To answer the question, I form a portfolio that is long stocks traded heavily outside the TSE and short those traded mostly on the TSE. Selecting the breakpoint for the portfolio formation is somewhat tricky. From Table 2.4 we see that portfolios balanced by asset count would be achieved by selecting the breakpoint to be 0.2 percent. Classifying stocks who have 0.3 percent of their volume traded outside of TSE as fundamentally different than those who have 0.1 percent of their volume traded outside of TSE seems to be a questionable approach. I choose a breakpoint of 50.0 percent, which results in unbalanced portfolios in terms of asset count but is more economically reasonable. After adjusting for size, the long portfolio contains 26 stocks and the short portfolio includes 272 stocks. For robustness, I consider various alternative specifications including breakpoints of 1, 5, 10, and 25 percent and removal of the size adjustment. The results, in terms of statistical significance, are robust to all the alternate specifications I consider.

Table 2.7 and Figure 2.5 provides the cumulative returns for the portfolios. Differences in the amount of volume cleared outside of TSE do not appear to associated with any shifts in relative capitalization after the market closure. The cumulative return of the long-short portfolio is not statically significant on any of the ten days following the event. Also, the cumulative returns are also not statistically different than a portfolio formed by randomly selecting 26 and 272 cross-listed stocks for purchase and sale. This finding coupled with that reported in the analysis of the binary classification of cross versus single-listed stocks suggests that it is the ability to trade outside of TSE and not the amount of trade cleared on OSE or NSE that is important to investors in the event of a market closure.

Did investors flock to the OSE or the NSE after the TSE failure? This may be answered by testing whether the percentage of trade cleared outside of TSE for cross-listed stocks increased after the market closure? OSE and NSE provide volume data at the monthly frequency. Hence, it is not possible to test for a change immediately after the January 18th event. I test whether the share of trade conducted outside of TSE was higher in February 2006 than in the fourth quarter of 2005. To be included in the sample, a stock must have volume data on at least one of the two outside exchanges in February 2006 and at least one
month in the final quarter of 2005; 776 stocks meet the criteria.

The mean share of outside trade in the fourth quarter of 2005 is 5.0 percent with standard error 0.6 percent. The mean percentage of outside trade increases to 7.1 percent in February 2006; its standard error is 0.8 percent. Through a paired difference in means t-test, I find that the increase is statistically significant with a t-statistic of 8.3. Of the 776 stocks under consideration, 610 had a greater share of their volume occur on OSE or NSE in February 2006 than in the fourth quarter of the previous year. The result suggests that (i) OSE and NSE have the capacity to absorb additional trade and (ii) investors shifted some of their trade away from the TSE after the unexpected market closure. Although the shift in trade may suggest a deterioration in investors’ confidence in the TSE, by easing the demands on the TSE, it also helps to reduce the likelihood of an additional failure.

**Liquidity Portfolios**

Amihud and Mendelson (1986) show that an asset’s expected return is a concave function of its illiquidity level. The concavity predicts that a systematic change in liquidity will asymmetrically affect stocks with different liquidity levels. Accordingly, a negative shock to systematic liquidity levels should lead to larger capital losses for more liquid stocks. This section tests the prediction by sorting stocks on a battery of liquidity-related characteristics, forming portfolios that are long the liquid stocks and short the illiquid stocks, and investigating the relative change in value of the liquidity portfolio subsequent to the market closure. To minimize contamination due to size and industry effects, the portfolios are industry-neutralized and size-adjusted as detailed earlier in the paper. For each sort, the tables and figures report and plot the results for the market adjusted and unadjusted models. The results are similar for both approaches; for brevity, I will limit my discussion to the unadjusted model.

When no direct measure of asset liquidity is available, researchers frequently employ asset turnover to identify liquid versus illiquid stocks. According to Amihud and Mendelson (1986), investors with a long investment horizon (low turnover investors) hold illiquid assets
and those with a short investment horizon (high turnover investors) hold liquid assets. If illiquidity represents the price impact of order flow (Kyle (1985), Black (1986), and Amihud (2002)), then high turnover suggests greater market depth, which in turn implies higher liquidity. I rank stocks according to their mean daily turnover over the period October 1, 2005 through December 31, 2005 and form a liquidity portfolio that is long 129 high turnover stocks and short 126 low turnover stocks. Table 2.8 reports and Figure 2.6 plots the cumulative returns for the portfolio for the market adjusted and unadjusted models. The results are consistent with the concavity of expected returns on liquidity. The liquidity portfolio has an immediate 5 percent capital loss on the day of the market closure. The capital loss is persistent and increases by an additional one percent over the ten day period. The result is statistically significant over the entire period and the portfolio’s return difference from that of a similarly structured randomly-formed portfolio is also statistically significant.

For robustness, I consider several additional measures of market depth. Table 2.10 and Figure 2.8 provides the cumulative returns for assets sorted by the average number of daily transactions. Table 2.12 and Figure 2.10 report the returns for assets sorted by their average daily monetary volume, which is defined as the sum of the price times volume over each transaction during the day. Table 2.9 and Figure 2.7 give the relative value changes for the liquidity portfolio formed by sorting stocks on their average turnover per trade, which is defined as the daily turnover divided by the number of transactions. Sorting by the number of daily transactions, the liquidity portfolio has an initial 3.7 percent capital loss that increases to 5.2 percent by the end of the month. For the liquidity portfolio formed after sorting by monetary volume, the initial relative loss in value is 4.1 percent. The loss increases to 6 percent over the ten day period. Approximately 2.8 percent of value is immediately lost for the portfolio formed after sorting by asset turnover per transaction. The capital loss increases to 3.8 percent by the end of the period. The returns are statistically significant and statistically different than those of a similarly structured portfolio with assets selected randomly.

The previous sorts and portfolio formations attempt to infer asset liquidity based on various
measurements of trading depth. Another approach is to proxy for a certain aspect of liquidity. The Amihud (2002) illiquidity measure, defined below, captures the price impact component of illiquidity:

\[
ILLIQ_i = \frac{1}{T} \sum_{t=1}^{T} \frac{|r_i^t|}{p_i^t \text{volume}_i^t},
\] (2.7)

where \(p_i^t\) is the price of asset \(i\) at time \(t\). Low absolute returns and high monetary volume are indicative of low price impact and high liquidity. High absolute returns and low monetary volume indicate high price impact and high illiquidity. I compute the illiquidity measure for each asset traded only on the TSE over the period October 1, 2005 through December 31, 2005. The liquidity portfolio is formed by purchasing the 126 assets with the lowest illiquidity and selling the 129 assets with the highest illiquidity. The portfolio is industry-neutral and size-controlled. Table 2.11 and Figure 2.9 provide the results for the liquidity portfolio based on the Amihud illiquidity sort. The performance of the liquidity portfolio over the ten day period is consistent with the concavity effect. Liquid stocks experience an initial and persistent 2.5 percent capital loss. The capital loss increases to almost 4 percent over the ten day period. The relative loss in value of the liquidity portfolio is statistically significant and the performance of the portfolio is statistically different than that of a randomly-formed portfolio.

It is interesting to investigate whether there is any correspondence between the assets inferred to be illiquid from the turnover sort and those that are estimated to be illiquid based on the Amihud (2002) illiquidity measure. I form a portfolio that is long assets identified to be liquid by the turnover criteria and the Amihud illiquidity measure and short assets identified to be illiquid through the two sorts. Because the portfolios are quintile ranked, if the two measures are independent, the long and short portfolios would each have approximately \(0.20 \times 126 = 25\) stocks. The double sorted portfolio is long and short, respectively, 80 and 82 stocks. Approximately 64 percent of stocks identified as liquid by the turnover sort are also categorized as liquid by the illiquidity measure, suggesting that the two criteria capture a common component of liquidity. The proportion is similar for the illiquid stocks. Table 2.14 and Figure 2.12 present the results for the liquidity portfolio. The cumulative returns for the double sort are nearly identical to those obtained by the original turnover
sorted portfolios.

2.3.4 Measured Liquidity Shifts

Does the traditional measure of illiquidity, the bid-ask spread, reflect the change in liquidity associated with the increased risk of unanticipated market closures subsequent to the January 18 TSE shutdown? If so, were changes in spreads persistent or did spreads revert to their prior levels?

With transactions data, I estimate the bid-ask spread for each stock traded on the TSE over three ten-day windows using the Gibbs sampler estimator suggested and implemented by Hasbrouck (2006). Details of the estimation procedure are included in the Appendix.

The first window is over the period beginning on December 30, 2005 and ending on January 17, 2006 and provides estimates of the bid-ask spread just prior to the market closure. The second window begins on January 18, 2006 and ends on January 31, 2006 and provides estimates of the bid-ask spread during and after the event. The final window, beginning on February 1, 2006 and ending on February 14, 2006 is included to test whether spreads reverted back to their prior levels.

To ensure an adequate sample size for the Bayesian estimation, stocks with less than 100 observations within each window are excluded. 2164 stocks meet the criteria, 1565 of which are stocks that trade only on the TSE. Further, any stock whose estimate within any of the three windows is in the 1 percent tail of estimates for that window is considered an outlier and is removed. 2078 stocks survive the filter, 1501 of which trade only on the TSE.

Table 2.15 reports the sample statistics for the estimated spreads across the three windows and the three categories. Panel A supplies statistics for the 2078 stocks traded on the TSE that meet the above criteria. The average spread across the three periods is approximately one quarter percent. The 80 percent confidence region is approximately 40 basis points. Panel B reports the sample statistics for the 1501 stocks that trade only on the TSE and Panel C presents the statistics for the 577 stocks that trade on at least one additional Japanese exchange. The reported statistics indicate that cross-listed stocks have spreads
approximately 5 basis points lower than those traded only on TSE, reflecting their higher liquidity.

Table 2.16 reports test statistics for systematic shifts in spreads across time periods. The first row reports the increase in average spreads from the ten-day period preceding the market closure to the ten-day period subsequent to the shutdown. Across all stocks traded on the TSE, the average bid-ask spread increased a statistically significant 2.8 basis points. The widening of spreads is larger for stocks that trade only the TSE (3.0 basis points) than those that trade on TSE and at least one additional Japanese exchange (2.2 basis points). On a relative basis, spreads increased slightly more than 10 percent, an economically significant shift. The second and third rows provide evidence that spreads quickly reverted to their levels prior to the market closure. The second row reports the change in spreads from the second to the third window. Overall, spreads narrowed by 2.6 basis points, a statistically significant result. The third row reports the difference in spreads from the third period to the first. We see that the differences are economically small and statistically insignificant, suggesting that spreads returned to their prior levels within approximately 20 trading days. Although the difference in spreads between the third and first period are statistically insignificant for the three categories, spreads of stocks traded only on the TSE widened with respect to those that are cross-listed by 1.14 basis points. The result is statistically significant with a standard error of $\sqrt{0.30^2 + 0.32^2} = 0.44$ basis points.

Overall, the results suggest that spreads capture some, but not all, of the increased illiquidity associated with potential market closures. Spreads widened immediately after the market closure reflecting the decreased expected illiquidity, but quickly returned to their prior levels long before the risk of additional unanticipated market closures had been dissolved. However, the spreads of stocks traded only on the TSE widened with respect to those that are cross-listed and the relative shift persisted throughout the estimation period; the result may be attributed to the increased sensitivity of stocks that trade only on the TSE to a TSE market closure.
2.4 Conclusion

This paper investigates the reaction of stocks traded on the Tokyo Stock Exchange to the unexpected market closure on January 18, 2006. On that day, the number of transactions, approximately 4.5 million, reached the capacity of the exchange’s computer systems and the market closed 20 minutes early. With the closure came news about the surprisingly limited capacity of the exchange, which limits its ability to handle a surge in trading activity. Investors were informed that exchange upgrades would likely take six months to a year.

Using transaction level data, I estimate the probability of at least one market closure over the subsequent six month period to be approximately 70 percent and the expected number of closures to be approximately 1.5. Hence, investors have reason to suspect that they may have difficulty trading in the near future. I compare the relative change in value of the 657 stocks that are only traded on the TSE to that of the 297 stocks that are also listed on either the Osaka Securities Exchange, or the Nagoya Stock Exchange, or both. I report that stocks listed only on the TSE experienced an approximate 2.2 percent capital loss relative to the stocks with an additional avenue for trade. I also compare the change in value of liquid stocks to illiquid stocks by sorting on a number of liquidity-related characteristics, such as asset turnover and the Amihud (2002) illiquidity measure. The results of the analysis are consistent with the concavity of expected returns in liquidity detailed by Amihud and Mendelson (1986). Each sort and portfolio formation confirms that liquid stocks lost value relative to illiquid stocks. For instance, over the ten day period subsequent to the market closure, high turnover stocks experienced a statistically significant 6.2 percent capital loss relative to low turnover stocks.

The relatively large value shifts after the January 18 market closure are surprising. There is no guarantee that the exchange will fail again. Even if the TSE reaches capacity and shuts down before its computer systems are sufficiently upgraded, an unanticipated market closure is likely to result in only a temporary loss in liquidity that automatically ‘resets’ the next day. It is not clear whether investors overreacted to the news or rationally updated their beliefs about the exchange’s ability to consistently provide liquidity. Perhaps the exchange’s failure
to provide adequate trading capacity is indicative of additional undiscovered problems. Regardless, the unexpected market closure coupled with the asymmetric changes in asset valuations provides additional evidence that investors place an economically significant value on the liquidity of their holdings.
Table 2.1: **Summary Statistics for the Exchanges.** This table reports the summary statistics for the Tokyo Stock Exchange, Nagoya Securities Exchange, and Osaka Stock Exchange. The first three lines are for stocks that are traded on a single exchange, with TSE on the first line, NSE on the second, and OSE on the third. The fourth line is for stocks that trade on the TSE and the NSE. The fifth line is for stocks that trade on the TSE and OSE. Statistics for stocks traded on the NSE and OSE are on the sixth line. The final line reports the statistics for stocks that trade on all three exchanges.

<table>
<thead>
<tr>
<th>Exchange</th>
<th>Number</th>
<th>Market Share by Value (percent)</th>
<th>Market Capitalization (¥1 Billion) Mean</th>
<th>Median</th>
<th>Trading Value (¥1 Million) Mean</th>
<th>Median</th>
<th>Annualized Turnover (percent) Mean</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>1552</td>
<td>34.6</td>
<td>115.9</td>
<td>31.1</td>
<td>756.6</td>
<td>135.6</td>
<td>224.8</td>
<td>92.8</td>
</tr>
<tr>
<td>N</td>
<td>99</td>
<td>0.3</td>
<td>15.3</td>
<td>8.8</td>
<td>35.8</td>
<td>3.7</td>
<td>76.4</td>
<td>11.7</td>
</tr>
<tr>
<td>O</td>
<td>367</td>
<td>1.4</td>
<td>19.7</td>
<td>9.5</td>
<td>131.6</td>
<td>21.4</td>
<td>135.6</td>
<td>62.4</td>
</tr>
<tr>
<td>TN</td>
<td>133</td>
<td>2.6</td>
<td>103.2</td>
<td>43.9</td>
<td>414.4</td>
<td>138.4</td>
<td>120.8</td>
<td>63.6</td>
</tr>
<tr>
<td>TO</td>
<td>543</td>
<td>23.9</td>
<td>228.8</td>
<td>67.3</td>
<td>1453.6</td>
<td>248.1</td>
<td>164.9</td>
<td>91.0</td>
</tr>
<tr>
<td>NO</td>
<td>11</td>
<td>0.1</td>
<td>29.2</td>
<td>9.0</td>
<td>32.5</td>
<td>20.9</td>
<td>66.5</td>
<td>34.6</td>
</tr>
<tr>
<td>TNO</td>
<td>142</td>
<td>37.2</td>
<td>1364.8</td>
<td>587.2</td>
<td>6768.5</td>
<td>3226.2</td>
<td>167.2</td>
<td>126.3</td>
</tr>
</tbody>
</table>
Table 2.2: **Industry Classification for TSE Only Stocks.** Industry classification for each stock listed on the TSE is obtained from the exchange’s website and is a snapshot taken on November 10, 2006. The industry classification of the 54 stocks that were traded on the TSE at the end of 2005, but were no longer listed on November 10, 2006 is obtained from Bloomberg.com. The table includes the 1552 stocks that are traded only on the TSE.

<table>
<thead>
<tr>
<th>Group</th>
<th>Industry</th>
<th>Count</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Transportation Equipment</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Transportation (Land)</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Transportation (Air)</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Transportation (Marine)</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Warehousing and Harbor Transportation Services</td>
<td>19</td>
<td>101</td>
</tr>
<tr>
<td>2</td>
<td>Banks</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Insurance</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Securities and Commodity Futures</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Real Estate</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Other Financing Business</td>
<td>30</td>
<td>139</td>
</tr>
<tr>
<td>3</td>
<td>Metal Products</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nonferrous Metals</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Iron</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Glass and Ceramic</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rubber</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Oil and Coal</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chemicals</td>
<td>94</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pulp and Paper</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Textiles</td>
<td>24</td>
<td>235</td>
</tr>
<tr>
<td>4</td>
<td>Wholesale Trade</td>
<td>129</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Retail Trade</td>
<td>148</td>
<td>277</td>
</tr>
<tr>
<td>5</td>
<td>Services</td>
<td>124</td>
<td>124</td>
</tr>
<tr>
<td>6</td>
<td>Foods</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Fishery, Agriculture, and Forestry</td>
<td>6</td>
<td>66</td>
</tr>
<tr>
<td>7</td>
<td>Construction</td>
<td>85</td>
<td>85</td>
</tr>
<tr>
<td>8</td>
<td>Machinery</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>9</td>
<td>Electric Appliances</td>
<td>142</td>
<td>142</td>
</tr>
<tr>
<td>10</td>
<td>Information and Communication</td>
<td>162</td>
<td>162</td>
</tr>
<tr>
<td>11</td>
<td>Power and Gas</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Precision Instruments</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pharmaceutical</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Other Products</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mining</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Exchange Traded Funds</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Real Estate Investment Trusts</td>
<td>26</td>
<td>141</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>1552</strong></td>
<td></td>
</tr>
</tbody>
</table>
Figure 2.1: Daily Equity Transactions. This figure plots the total number of equity transactions processed by the Tokyo Stock Exchange on each day over the period October 1, 2005 through February 28, 2006. The number of transactions are in millions.
Table 2.3: Cumulative Returns by Industry Group. An equal-weight portfolio is formed for each of the 11 industry groups identified by Table 2.2 using all stocks listed on the Tokyo Stock Exchange. The cumulative simple return over the $d$-day holding period with portfolio formation occurring on January 18, 2006 is denoted by $R_d = \exp(\alpha_1 + \cdots + \alpha_d) - 1$ where the alphas are the coefficients from the following regression, $r_t = \alpha + \alpha_1 \psi_1 + \cdots + \alpha_d \psi_d + \epsilon_t$, estimated over the period October 1, 2005 through February 28, 2006. The name of each industry group is presented next to the group number in parentheses. Below the industry name is the number of stocks contained in the respective group. Returns are reported as a percentage and standard errors are in parenthesis below their respective returns.

<table>
<thead>
<tr>
<th>Num</th>
<th>Industry Group</th>
<th>$R_1$</th>
<th>$R_2$</th>
<th>$R_3$</th>
<th>$R_4$</th>
<th>$R_5$</th>
<th>$R_6$</th>
<th>$R_7$</th>
<th>$R_8$</th>
<th>$R_9$</th>
<th>$R_{10}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Transportation</td>
<td>-5.42</td>
<td>-2.49</td>
<td>-2.54</td>
<td>-5.12</td>
<td>-3.80</td>
<td>-3.38</td>
<td>-2.67</td>
<td>-1.02</td>
<td>0.38</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.29)</td>
<td>(0.23)</td>
<td>(0.30)</td>
<td>(0.41)</td>
<td>(0.43)</td>
<td>(0.43)</td>
<td>(0.49)</td>
<td>(0.53)</td>
<td>(0.52)</td>
<td>(0.54)</td>
</tr>
<tr>
<td>2</td>
<td>Finance</td>
<td>-5.72</td>
<td>-1.28</td>
<td>-2.53</td>
<td>-6.24</td>
<td>-3.58</td>
<td>-2.44</td>
<td>-0.61</td>
<td>1.59</td>
<td>1.64</td>
<td>1.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.29)</td>
<td>(0.21)</td>
<td>(0.30)</td>
<td>(0.43)</td>
<td>(0.35)</td>
<td>(0.32)</td>
<td>(0.43)</td>
<td>(0.47)</td>
<td>(0.50)</td>
<td>(0.51)</td>
</tr>
<tr>
<td>3</td>
<td>Industrial</td>
<td>-6.00</td>
<td>-2.50</td>
<td>-2.53</td>
<td>-5.46</td>
<td>-3.87</td>
<td>-2.99</td>
<td>-1.89</td>
<td>-0.36</td>
<td>0.47</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.17)</td>
<td>(0.13)</td>
<td>(0.17)</td>
<td>(0.22)</td>
<td>(0.22)</td>
<td>(0.22)</td>
<td>(0.30)</td>
<td>(0.32)</td>
<td>(0.34)</td>
<td>(0.36)</td>
</tr>
<tr>
<td>4</td>
<td>Wholesale/Retail</td>
<td>-6.01</td>
<td>-2.32</td>
<td>-2.97</td>
<td>-6.10</td>
<td>-4.31</td>
<td>-3.80</td>
<td>-3.08</td>
<td>-1.59</td>
<td>-0.75</td>
<td>-0.83</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.21)</td>
<td>(0.23)</td>
<td>(0.26)</td>
<td>(0.30)</td>
<td>(0.27)</td>
<td>(0.28)</td>
<td>(0.34)</td>
<td>(0.36)</td>
<td>(0.40)</td>
<td>(0.41)</td>
</tr>
<tr>
<td>5</td>
<td>Services</td>
<td>-7.82</td>
<td>-3.18</td>
<td>-5.07</td>
<td>-9.34</td>
<td>-6.01</td>
<td>-5.17</td>
<td>-4.88</td>
<td>-3.56</td>
<td>-3.10</td>
<td>-3.36</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.37)</td>
<td>(0.34)</td>
<td>(0.47)</td>
<td>(0.65)</td>
<td>(0.51)</td>
<td>(0.48)</td>
<td>(0.73)</td>
<td>(0.77)</td>
<td>(0.82)</td>
<td>(0.91)</td>
</tr>
<tr>
<td>6</td>
<td>Food/Agriculture</td>
<td>-3.64</td>
<td>-1.15</td>
<td>-1.20</td>
<td>-2.91</td>
<td>-2.15</td>
<td>-2.22</td>
<td>-1.40</td>
<td>-0.42</td>
<td>0.37</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.28)</td>
<td>(0.21)</td>
<td>(0.27)</td>
<td>(0.32)</td>
<td>(0.31)</td>
<td>(0.32)</td>
<td>(0.35)</td>
<td>(0.35)</td>
<td>(0.40)</td>
<td>(0.37)</td>
</tr>
<tr>
<td>7</td>
<td>Construction</td>
<td>-5.38</td>
<td>-1.69</td>
<td>-2.12</td>
<td>-4.58</td>
<td>-3.13</td>
<td>-2.35</td>
<td>-0.71</td>
<td>1.14</td>
<td>2.46</td>
<td>3.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.30)</td>
<td>(0.26)</td>
<td>(0.45)</td>
<td>(0.58)</td>
<td>(0.48)</td>
<td>(0.51)</td>
<td>(0.56)</td>
<td>(0.56)</td>
<td>(0.73)</td>
<td>(0.78)</td>
</tr>
<tr>
<td>8</td>
<td>Machinery</td>
<td>-6.48</td>
<td>-2.22</td>
<td>-2.76</td>
<td>-6.15</td>
<td>-4.70</td>
<td>-3.80</td>
<td>-2.57</td>
<td>-0.49</td>
<td>0.77</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.27)</td>
<td>(0.19)</td>
<td>(0.25)</td>
<td>(0.29)</td>
<td>(0.29)</td>
<td>(0.34)</td>
<td>(0.40)</td>
<td>(0.41)</td>
<td>(0.48)</td>
<td>(0.48)</td>
</tr>
<tr>
<td>9</td>
<td>Electric Appliances</td>
<td>-5.91</td>
<td>-2.42</td>
<td>-2.46</td>
<td>-6.19</td>
<td>-4.51</td>
<td>-3.42</td>
<td>-2.26</td>
<td>0.21</td>
<td>0.51</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.26)</td>
<td>(0.20)</td>
<td>(0.26)</td>
<td>(0.31)</td>
<td>(0.31)</td>
<td>(0.34)</td>
<td>(0.37)</td>
<td>(0.42)</td>
<td>(0.44)</td>
<td>(0.47)</td>
</tr>
<tr>
<td>10</td>
<td>Information/Communication</td>
<td>-7.90</td>
<td>-3.51</td>
<td>-6.27</td>
<td>-11.73</td>
<td>-7.56</td>
<td>-6.57</td>
<td>-6.11</td>
<td>-4.71</td>
<td>-4.97</td>
<td>-5.45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.43)</td>
<td>(0.50)</td>
<td>(0.65)</td>
<td>(0.83)</td>
<td>(0.79)</td>
<td>(0.77)</td>
<td>(1.00)</td>
<td>(1.05)</td>
<td>(1.10)</td>
<td>(1.15)</td>
</tr>
<tr>
<td>11</td>
<td>Other</td>
<td>-4.34</td>
<td>-1.59</td>
<td>-1.64</td>
<td>-3.68</td>
<td>-2.34</td>
<td>-2.22</td>
<td>-1.54</td>
<td>0.02</td>
<td>0.17</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.25)</td>
<td>(0.17)</td>
<td>(0.23)</td>
<td>(0.36)</td>
<td>(0.29)</td>
<td>(0.29)</td>
<td>(0.33)</td>
<td>(0.35)</td>
<td>(0.38)</td>
<td>(0.40)</td>
</tr>
</tbody>
</table>
Figure 2.2: Simulated Market Shutdowns. This figure plots the histogram for the total number of times the Tokyo Stock Exchange is predicted to shutdown during the six month period subsequent to the January 18, 2006 market closure. Two million simulations of the following AR(1) specification for total daily equity transactions, in millions, were generated: \( \tau_{t+1} = 0.5065 + 0.6363 \tau_t + \epsilon_{t+1} \). The parameters of the model are the OLS estimates obtained over the sample period October 1, 2005 through January 18, 2006 and the shocks in the simulations are the residuals from the regression, resampled with replacement. Both signs of each residual are included. A shutdown occurs when the number of simulated equity transactions on a given day exceeds 2.03 million, the number of equity transactions that occurred on January 18, 2006.
Table 2.4: Cumulative Density Function for Percentage of Trade Transacted Outside of TSE for Cross-Listed Stocks. The following table reports the CDF for the percentage of total share volume cleared by the OSE and NSE in the fourth quarter of 2005 for the 818 stocks listed on TSE and at least one additional exchange.

<table>
<thead>
<tr>
<th>Outside Trade</th>
<th>Count</th>
<th>CDF</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>62</td>
<td>7.6</td>
</tr>
<tr>
<td>0.1</td>
<td>315</td>
<td>38.5</td>
</tr>
<tr>
<td>0.2</td>
<td>414</td>
<td>50.0</td>
</tr>
<tr>
<td>0.5</td>
<td>559</td>
<td>68.3</td>
</tr>
<tr>
<td>1.0</td>
<td>655</td>
<td>80.1</td>
</tr>
<tr>
<td>10.0</td>
<td>752</td>
<td>91.9</td>
</tr>
<tr>
<td>50.0</td>
<td>760</td>
<td>92.9</td>
</tr>
<tr>
<td>75.0</td>
<td>765</td>
<td>93.5</td>
</tr>
<tr>
<td>90.0</td>
<td>802</td>
<td>98.0</td>
</tr>
<tr>
<td>95.0</td>
<td>817</td>
<td>99.9</td>
</tr>
<tr>
<td>100.0</td>
<td>818</td>
<td>100.0</td>
</tr>
</tbody>
</table>
Table 2.5: Portfolios Formed by Market Capitalization. All equities listed on the Tokyo Stock Exchange are sorted by their average daily market capitalization over the period October 1, 2005 through December 31, 2005 within each of the 11 industry groups identified by Table 2.2. A zero-cost portfolio is formed by purchasing the quintile within each industry group with the highest market capitalization and selling the quintile within each industry group with the lowest market capitalization. The cumulative simple return over the d-day holding period with portfolio formation occurring on January 18, 2006 is denoted by \( R_t \equiv \exp(\alpha_1 + \cdots + \alpha_d) - 1 \) where the alphas are the coefficients from the following regression, \( r_t = \alpha + \beta r^{m}_t + \alpha_1 \psi_1 + \cdots + \alpha_{10} \psi_{10} + \epsilon_t \), estimated over the period October 1, 2005 through February 28, 2006 where \( r^{m}_t \) is the return on the Nikkei 225 Index. The first three lines set \( \psi = 1 \) and are adjusted for a market factor. The following three lines set \( \psi = 0 \) and are unadjusted. Cumulative returns and \( \alpha \) are reported in percent. Standard errors are in parenthesis.

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>( \alpha_1 )</th>
<th>( \alpha_2 )</th>
<th>( \alpha_3 )</th>
<th>( \alpha_4 )</th>
<th>( \alpha_5 )</th>
<th>( \alpha_6 )</th>
<th>( \alpha_7 )</th>
<th>( \alpha_8 )</th>
<th>( \alpha_9 )</th>
<th>( \alpha_{10} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big</td>
<td>-0.02</td>
<td>0.80</td>
<td>-2.30</td>
<td>-0.45</td>
<td>-0.73</td>
<td>-2.22</td>
<td>-1.07</td>
<td>-0.49</td>
<td>-0.36</td>
<td>-0.98</td>
<td>-0.87</td>
<td>-1.23</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.03)</td>
<td>(0.22)</td>
<td>(0.17)</td>
<td>(0.24)</td>
<td>(0.31)</td>
<td>(0.27)</td>
<td>(0.29)</td>
<td>(0.31)</td>
<td>(0.35)</td>
<td>(0.39)</td>
<td>(0.42)</td>
</tr>
<tr>
<td>Small</td>
<td>-0.06</td>
<td>0.53</td>
<td>-6.23</td>
<td>-2.92</td>
<td>-3.75</td>
<td>-6.59</td>
<td>-5.27</td>
<td>-4.38</td>
<td>-4.34</td>
<td>-4.60</td>
<td>-4.08</td>
<td>-4.27</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.03)</td>
<td>(0.22)</td>
<td>(0.16)</td>
<td>(0.23)</td>
<td>(0.32)</td>
<td>(0.29)</td>
<td>(0.29)</td>
<td>(0.35)</td>
<td>(0.39)</td>
<td>(0.41)</td>
<td>(0.44)</td>
</tr>
<tr>
<td>Big - Small</td>
<td>0.04</td>
<td>0.26</td>
<td>3.94</td>
<td>2.47</td>
<td>3.02</td>
<td>4.37</td>
<td>4.20</td>
<td>3.89</td>
<td>3.98</td>
<td>3.62</td>
<td>3.21</td>
<td>3.03</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.31)</td>
<td>(0.31)</td>
<td>(0.33)</td>
<td>(0.45)</td>
<td>(0.49)</td>
<td>(0.41)</td>
<td>(0.47)</td>
<td>(0.52)</td>
<td>(0.56)</td>
<td>(0.60)</td>
</tr>
<tr>
<td>Big</td>
<td>0.10</td>
<td>-4.68</td>
<td>-1.22</td>
<td>-1.60</td>
<td>-4.82</td>
<td>-2.40</td>
<td>-1.93</td>
<td>-0.73</td>
<td>1.32</td>
<td>1.76</td>
<td>1.75</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.23)</td>
<td>(0.17)</td>
<td>(0.24)</td>
<td>(0.32)</td>
<td>(0.27)</td>
<td>(0.29)</td>
<td>(0.31)</td>
<td>(0.35)</td>
<td>(0.39)</td>
<td>(0.41)</td>
<td></td>
</tr>
<tr>
<td>Small</td>
<td>0.01</td>
<td>-7.74</td>
<td>-3.42</td>
<td>-4.31</td>
<td>-8.22</td>
<td>-6.11</td>
<td>-5.30</td>
<td>-4.58</td>
<td>-3.16</td>
<td>-2.43</td>
<td>-2.40</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.24)</td>
<td>(0.16)</td>
<td>(0.23)</td>
<td>(0.34)</td>
<td>(0.29)</td>
<td>(0.29)</td>
<td>(0.35)</td>
<td>(0.39)</td>
<td>(0.41)</td>
<td>(0.43)</td>
<td></td>
</tr>
<tr>
<td>Big - Small</td>
<td>0.09</td>
<td>3.06</td>
<td>2.20</td>
<td>2.71</td>
<td>3.39</td>
<td>3.71</td>
<td>3.36</td>
<td>3.86</td>
<td>4.48</td>
<td>4.19</td>
<td>4.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.33)</td>
<td>(0.23)</td>
<td>(0.34)</td>
<td>(0.47)</td>
<td>(0.40)</td>
<td>(0.41)</td>
<td>(0.47)</td>
<td>(0.52)</td>
<td>(0.56)</td>
<td>(0.60)</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2.3: Portfolios Formed by Market Capitalization. The two figures plot the simple cumulative returns reported by Table 2.5 for the zero-cost industry-neutral portfolio that is long the quintile of stocks with the highest market capitalization and short the quintile of stocks with the lowest market capitalization. The cumulative returns from the market adjusted model is on the left and from the unadjusted model is on the right. The solid line is the cumulative return of the zero-cost portfolio in percent. The dashed lines are the 95 percent confidence bands for the portfolio’s estimated return. The dotted lines are the 95 percent confidence bands for the cumulative return of an industry-neutral and size-adjusted portfolio that is long and short, respectively, 309 and 307 randomly selected stocks.
Table 2.6: Portfolios Formed by Exchange. All equities listed on the Tokyo Stock Exchange (TSE) are sorted by size within their respective industry groups. The smallest 30 percent of stocks and the largest 30 percent of firms within each industry group are removed to control for size effects. A zero-cost portfolio is formed by purchasing all assets that trade only on the TSE and selling all assets that are cross-listed on the TSE and at least one additional Japanese exchange. The long and short components of the portfolio include, respectively, 657 and 297 stocks. The cumulative simple return over the $d$-day holding period with portfolio formation occurring on January 18, 2006 is denoted by $R_d \equiv \exp(\alpha_1 + \cdots + \alpha_d) - 1$ where the alphas are the coefficients from the following regression, $r_t = \alpha + \beta r_{t-1}^m + \alpha_1 \psi_1 + \cdots + \alpha_{10} \psi_{10} + \varepsilon_t$, estimated over the period October 1, 2005 through February 28, 2006 where $r_{t-1}^m$ is the return on the Nikkei 225 Index. The first three lines set $I = 1$ and are adjusted for a market factor. The following three lines set $I = 0$ and are unadjusted. Cumulative returns and $\alpha$ are reported in percent. Standard errors are in parenthesis.

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$\alpha_1$</th>
<th>$\alpha_2$</th>
<th>$\alpha_3$</th>
<th>$\alpha_4$</th>
<th>$\alpha_5$</th>
<th>$\alpha_6$</th>
<th>$\alpha_7$</th>
<th>$\alpha_8$</th>
<th>$\alpha_9$</th>
<th>$\alpha_{10}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSE Only</td>
<td>0.02</td>
<td>0.79</td>
<td>-4.11</td>
<td>-1.80</td>
<td>-2.62</td>
<td>-4.51</td>
<td>-3.42</td>
<td>-2.47</td>
<td>-2.71</td>
<td>-3.46</td>
<td>-3.31</td>
<td>-3.65</td>
</tr>
<tr>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.15)</td>
<td>(0.18)</td>
<td>(0.23)</td>
<td>(0.29)</td>
<td>(0.27)</td>
<td>(0.27)</td>
<td>(0.27)</td>
<td>(0.37)</td>
<td>(0.39)</td>
<td>(0.41)</td>
<td>(0.44)</td>
</tr>
<tr>
<td>Cross-Listed</td>
<td>0.02</td>
<td>0.70</td>
<td>-3.57</td>
<td>-1.10</td>
<td>-1.56</td>
<td>-2.91</td>
<td>-2.49</td>
<td>-1.63</td>
<td>-1.14</td>
<td>-1.61</td>
<td>-1.19</td>
<td>-1.53</td>
</tr>
<tr>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.19)</td>
<td>(0.16)</td>
<td>(0.19)</td>
<td>(0.23)</td>
<td>(0.24)</td>
<td>(0.27)</td>
<td>(0.32)</td>
<td>(0.33)</td>
<td>(0.35)</td>
<td>(0.37)</td>
<td></td>
</tr>
<tr>
<td>Long-Short Portfolio</td>
<td>-0.01</td>
<td>0.00</td>
<td>-0.53</td>
<td>-0.70</td>
<td>-1.07</td>
<td>-1.60</td>
<td>-0.94</td>
<td>-0.83</td>
<td>-1.57</td>
<td>-1.85</td>
<td>-2.12</td>
<td>-2.12</td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.03)</td>
<td>(0.26)</td>
<td>(0.29)</td>
<td>(0.36)</td>
<td>(0.46)</td>
<td>(0.43)</td>
<td>(0.44)</td>
<td>(0.58)</td>
<td>(0.62)</td>
<td>(0.65)</td>
<td>(0.70)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$\alpha_1$</th>
<th>$\alpha_2$</th>
<th>$\alpha_3$</th>
<th>$\alpha_4$</th>
<th>$\alpha_5$</th>
<th>$\alpha_6$</th>
<th>$\alpha_7$</th>
<th>$\alpha_8$</th>
<th>$\alpha_9$</th>
<th>$\alpha_{10}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSE Only</td>
<td>0.12</td>
<td>-6.16</td>
<td>-2.46</td>
<td>-3.37</td>
<td>-6.71</td>
<td>-4.55</td>
<td>-3.69</td>
<td>-3.01</td>
<td>-1.48</td>
<td>-1.10</td>
<td>-1.15</td>
<td></td>
</tr>
<tr>
<td>(0.01)</td>
<td>(0.17)</td>
<td>(0.23)</td>
<td>(0.30)</td>
<td>(0.27)</td>
<td>(0.36)</td>
<td>(0.38)</td>
<td>(0.40)</td>
<td>(0.43)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cross-Listed</td>
<td>0.13</td>
<td>-5.65</td>
<td>-1.78</td>
<td>-2.33</td>
<td>-5.23</td>
<td>-3.66</td>
<td>-2.91</td>
<td>-1.48</td>
<td>0.40</td>
<td>1.12</td>
<td>1.08</td>
<td></td>
</tr>
<tr>
<td>(0.01)</td>
<td>(0.20)</td>
<td>(0.19)</td>
<td>(0.24)</td>
<td>(0.24)</td>
<td>(0.26)</td>
<td>(0.32)</td>
<td>(0.34)</td>
<td>(0.37)</td>
<td>(0.38)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long-Short Portfolio</td>
<td>-0.01</td>
<td>-0.50</td>
<td>-0.69</td>
<td>-1.04</td>
<td>-1.54</td>
<td>-0.89</td>
<td>-0.78</td>
<td>-1.53</td>
<td>-1.91</td>
<td>-2.21</td>
<td>-2.23</td>
<td></td>
</tr>
<tr>
<td>(0.02)</td>
<td>(0.29)</td>
<td>(0.37)</td>
<td>(0.48)</td>
<td>(0.44)</td>
<td>(0.44)</td>
<td>(0.58)</td>
<td>(0.61)</td>
<td>(0.65)</td>
<td>(0.69)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2.4: Portfolios Formed by Exchange. The two figures plot the simple cumulative returns reported by Table 2.6 for the zero-cost industry-neutral portfolio that is long the stocks traded only on the TSE and short the stocks that are cross-listed on the TSE and at least one additional Japanese exchange. The cumulative returns from the market adjusted model is on the left and from the unadjusted model is on the right. The solid line is the cumulative return of the zero-cost portfolio in percent. The dashed lines are the 95 percent confidence bands for the portfolio’s estimated return. The dotted lines are the 95 percent confidence bands for the cumulative return of an industry-neutral and size-adjusted portfolio that is long and short, respectively, 309 and 307 randomly selected stocks.
Table 2.7: Portfolios Formed by Relative Volume Outside of TSE. All equities listed on the TSE and at least one additional Japanese exchange are sorted by the percentage of volume cleared on the TSE. The smallest 30 percent of stocks and the largest 30 percent of firms within each industry group are removed to control for size effects. A zero-cost portfolio is formed by purchasing assets with less than half of trade volume transacted on TSE and selling the remaining assets. The long and short components of the portfolio include, respectively, 26 and 272 stocks. The cumulative simple return over the \( d \)-day holding period with portfolio formation occurring on January 18, 2006 is denoted by \( R_d \equiv \exp(\alpha_1 + \cdots + \alpha_d) - 1 \) where the alphas are the coefficients from the following regression, \( r_t = \alpha + \beta r_t^{10} + \alpha_1 \psi_1 + \cdots + \alpha_{10} \psi_{10} + \varepsilon_t \), estimated over the period October 1, 2005 through February 28, 2006 where \( r_t^{10} \) is the return on the Nikkei 225 Index. The first three lines set \( I = 1 \) and are adjusted for a market factor. The following three lines set \( I = 0 \) and are unadjusted. Cumulative returns and \( \alpha \) are reported in percent. Standard errors are in parenthesis.

<table>
<thead>
<tr>
<th>Parameter Estimates</th>
<th>Portfolio</th>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>( \alpha_1 )</th>
<th>( \alpha_2 )</th>
<th>( \alpha_3 )</th>
<th>( \alpha_4 )</th>
<th>( \alpha_5 )</th>
<th>( \alpha_6 )</th>
<th>( \alpha_7 )</th>
<th>( \alpha_8 )</th>
<th>( \alpha_9 )</th>
<th>( \alpha_{10} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Outside</td>
<td>-0.03</td>
<td>0.58</td>
<td>-4.15</td>
<td>-1.56</td>
<td>-1.40</td>
<td>-2.66</td>
<td>-3.24</td>
<td>-2.88</td>
<td>-1.69</td>
<td>-1.60</td>
<td>-1.45</td>
<td>-1.78</td>
<td></td>
</tr>
<tr>
<td>Volume</td>
<td>(0.05)</td>
<td>(0.07)</td>
<td>(0.61)</td>
<td>(0.57)</td>
<td>(0.65)</td>
<td>(0.83)</td>
<td>(1.07)</td>
<td>(1.16)</td>
<td>(1.20)</td>
<td>(1.20)</td>
<td>(1.09)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Outside</td>
<td>0.03</td>
<td>0.71</td>
<td>-3.52</td>
<td>-1.05</td>
<td>-1.57</td>
<td>-2.93</td>
<td>-2.41</td>
<td>-1.52</td>
<td>-1.09</td>
<td>-1.61</td>
<td>-1.16</td>
<td>-1.51</td>
<td></td>
</tr>
<tr>
<td>Volume</td>
<td>(0.01)</td>
<td>(0.02)</td>
<td>(0.21)</td>
<td>(0.17)</td>
<td>(0.20)</td>
<td>(0.24)</td>
<td>(0.24)</td>
<td>(0.27)</td>
<td>(0.33)</td>
<td>(0.35)</td>
<td>(0.37)</td>
<td>(0.39)</td>
<td></td>
</tr>
<tr>
<td>Long-Short</td>
<td>-0.06</td>
<td>-0.13</td>
<td>-0.64</td>
<td>-0.51</td>
<td>0.18</td>
<td>0.27</td>
<td>-0.83</td>
<td>-1.36</td>
<td>-0.60</td>
<td>0.01</td>
<td>-0.28</td>
<td>-0.27</td>
<td></td>
</tr>
<tr>
<td>Portfolio</td>
<td>(0.04)</td>
<td>(0.08)</td>
<td>(0.69)</td>
<td>(0.57)</td>
<td>(0.69)</td>
<td>(0.83)</td>
<td>(0.86)</td>
<td>(0.95)</td>
<td>(1.14)</td>
<td>(1.17)</td>
<td>(1.25)</td>
<td>(1.30)</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2.5: Portfolios Formed by Relative Volume Outside of TSE. The two figures plot the simple cumulative returns reported by Table 2.7 for the zero-cost size-adjusted portfolio that is long the cross-listed stocks with less than half of trade volume occurring on TSE and short the remaining cross-listed stocks. The cumulative returns from the market adjusted model is on the left and from the unadjusted model is on the right. The solid line is the cumulative return of the zero-cost portfolio in percent. The dashed lines are the 95 percent confidence bands for the portfolio’s estimated return. The dotted lines are the 95 percent confidence bands for the cumulative return of an industry-neutral and size-adjusted portfolio that is long and short, respectively, 26 and 272 randomly selected cross-listed stocks.
Table 2.8: Portfolios Formed by Share Turnover. All equities listed on the Tokyo Stock Exchange are sorted by size within their respective industry groups. The smallest 30 percent of stocks and the largest 30 percent of firms within each industry group are removed to control for size effects. After removing stocks that are traded on multiple exchanges in Japan, the remaining assets are sorted by their average daily turnover over the period October 1, 2005 through December 31, 2005 within each of the 11 industry groups identified by Table 2.2. A zero-cost portfolio is formed by purchasing the quintile within each industry group with the highest turnover and selling the quintile within each industry group with the lowest turnover. The long and short components of the portfolio include, respectively, 129 and 126 stocks. The cumulative simple return over the $d$-day holding period with portfolio formation occurring on January 18, 2006 is denoted by $R_d = \exp(\alpha + \cdots + \alpha_d) - 1$ where the alphas are the coefficients from the following regression, $r_t = \alpha + \beta r_{m,t} + \alpha_1 \psi_1 + \cdots + \alpha_{10} \psi_{10} + \epsilon_t$, estimated over the period October 1, 2005 through February 28, 2006 where $r_{m,t}$ is the return on the Nikkei 225 Index. The first three lines set $I = 1$ and are unadjusted. Cumulative returns and $\alpha$ are reported in percent. Standard errors are in parenthesis.

<table>
<thead>
<tr>
<th>Parameter Estimates</th>
<th>Portfolio</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$\alpha_1$</th>
<th>$\alpha_2$</th>
<th>$\alpha_3$</th>
<th>$\alpha_4$</th>
<th>$\alpha_5$</th>
<th>$\alpha_6$</th>
<th>$\alpha_7$</th>
<th>$\alpha_8$</th>
<th>$\alpha_9$</th>
<th>$\alpha_{10}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liquid</td>
<td>0.07</td>
<td>1.10</td>
<td>-6.11</td>
<td>-3.54</td>
<td>-5.85</td>
<td>-9.83</td>
<td>-7.56</td>
<td>-5.50</td>
<td>-6.97</td>
<td>-8.82</td>
<td>-9.03</td>
<td>-9.92</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.06)</td>
<td>(0.38)</td>
<td>(0.68)</td>
<td>(0.87)</td>
<td>(0.97)</td>
<td>(0.95)</td>
<td>(1.41)</td>
<td>(1.49)</td>
<td>(1.59)</td>
<td>(1.68)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illiquid</td>
<td>0.04</td>
<td>0.39</td>
<td>-2.96</td>
<td>-1.05</td>
<td>-1.25</td>
<td>-1.74</td>
<td>-1.44</td>
<td>-1.96</td>
<td>-1.95</td>
<td>-1.72</td>
<td>-1.51</td>
<td>-1.47</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.02)</td>
<td>(0.29)</td>
<td>(0.29)</td>
<td>(0.43)</td>
<td>(0.43)</td>
<td>(0.46)</td>
<td>(0.47)</td>
<td>(0.45)</td>
<td>(0.47)</td>
<td>(0.55)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long-Short</td>
<td>0.30</td>
<td>0.71</td>
<td>-3.14</td>
<td>-2.48</td>
<td>-4.60</td>
<td>-8.08</td>
<td>-6.12</td>
<td>-3.53</td>
<td>-5.01</td>
<td>-7.10</td>
<td>-7.52</td>
<td>-8.34</td>
<td></td>
</tr>
<tr>
<td>Portfolio</td>
<td>(0.06)</td>
<td>(0.06)</td>
<td>(0.47)</td>
<td>(0.72)</td>
<td>(0.92)</td>
<td>(1.12)</td>
<td>(1.07)</td>
<td>(1.07)</td>
<td>(1.50)</td>
<td>(1.57)</td>
<td>(1.68)</td>
<td>(1.78)</td>
<td></td>
</tr>
<tr>
<td>Liquid</td>
<td>0.21</td>
<td>-9.22</td>
<td>-4.52</td>
<td>-6.93</td>
<td>-12.97</td>
<td>-9.15</td>
<td>-7.25</td>
<td>-7.32</td>
<td>-5.91</td>
<td>-5.74</td>
<td>-6.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.43)</td>
<td>(0.68)</td>
<td>(0.87)</td>
<td>(1.06)</td>
<td>(0.97)</td>
<td>(0.95)</td>
<td>(1.40)</td>
<td>(1.48)</td>
<td>(1.58)</td>
<td>(1.64)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illiquid</td>
<td>0.10</td>
<td>-4.15</td>
<td>-1.44</td>
<td>-1.69</td>
<td>-3.07</td>
<td>-2.11</td>
<td>-2.68</td>
<td>-2.14</td>
<td>-0.60</td>
<td>-0.23</td>
<td>-0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.29)</td>
<td>(0.29)</td>
<td>(0.29)</td>
<td>(0.42)</td>
<td>(0.42)</td>
<td>(0.46)</td>
<td>(0.47)</td>
<td>(0.46)</td>
<td>(0.49)</td>
<td>(0.56)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long-Short</td>
<td>0.11</td>
<td>-5.07</td>
<td>-3.08</td>
<td>-5.24</td>
<td>-9.90</td>
<td>-7.04</td>
<td>-4.57</td>
<td>-5.17</td>
<td>-5.31</td>
<td>-5.52</td>
<td>-6.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Portfolio</td>
<td>(0.05)</td>
<td>(0.52)</td>
<td>(0.73)</td>
<td>(0.93)</td>
<td>(1.15)</td>
<td>(1.07)</td>
<td>(1.07)</td>
<td>(1.49)</td>
<td>(1.56)</td>
<td>(1.67)</td>
<td>(1.75)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2.6: Portfolios Formed by Mean Share Turnover. The two figures plot the simple cumulative returns reported by Table 2.8 for the zero-cost industry-neutral portfolio that is long the quintile of stocks with the highest turnover and short the quintile of stocks with the lowest turnover. The cumulative returns from the market adjusted model is on the left and from the unadjusted model is on the right. The solid line is the cumulative return of the zero-cost portfolio in percent. The dashed lines are the 95 percent confidence bands for the portfolio’s estimated return. The dotted lines are the 95 percent confidence bands for the cumulative return of an industry-neutral and size-adjusted portfolio that is long and short, respectively, 129 and 126 randomly selected stocks.

Market Adjusted

Not Market Adjusted
Table 2.9: Portfolios Formed by Turnover per Trade. All equities listed on the Tokyo Stock Exchange are sorted by size within their respective industry groups. The smallest 30 percent of stocks and the largest 30 percent of firms within each industry group are removed to control for size effects. After removing stocks that are traded on multiple exchanges in Japan, the remaining assets are sorted by their average turnover per transaction over the period October 1, 2005 through December 31, 2005 within each of the 11 industry groups identified by Table 2.2. A zero-cost portfolio is formed by purchasing the quintile within each industry group with the highest turnover per transaction and selling the quintile within each industry group with the lowest turnover per transaction. The long and short components of the portfolio include, respectively, 129 and 126 stocks. The cumulative simple return over the period October 1, 2005 through December 31, 2005 within each of the 11 industry groups is denoted by $R_{t}$ and is estimated over the period October 1, 2005 through February 28, 2006 where $r_{t}^{m}$ is the return on the Nikkei 225 Index. The first three lines set $I = 1$ and are adjusted for a market factor. The following three lines set $I = 0$ and are unadjusted. Cumulative returns and $\alpha$ are reported in percent. Standard errors are in parenthesis.

<table>
<thead>
<tr>
<th>Parameter Estimates</th>
<th>Portfolio</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$\alpha_{1}$</th>
<th>$\alpha_{2}$</th>
<th>$\alpha_{3}$</th>
<th>$\alpha_{4}$</th>
<th>$\alpha_{5}$</th>
<th>$\alpha_{6}$</th>
<th>$\alpha_{7}$</th>
<th>$\alpha_{8}$</th>
<th>$\alpha_{9}$</th>
<th>$\alpha_{10}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Liquid</td>
<td>0.10</td>
<td>0.82</td>
<td>-5.47</td>
<td>-2.71</td>
<td>-4.19</td>
<td>-6.96</td>
<td>-5.41</td>
<td>-4.13</td>
<td>-5.05</td>
<td>-6.17</td>
<td>-6.59</td>
<td>-6.98</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.05)</td>
<td>(0.07)</td>
<td>(0.37)</td>
<td>(0.40)</td>
<td>(0.56)</td>
<td>(0.78)</td>
<td>(0.67)</td>
<td>(0.56)</td>
<td>(1.02)</td>
<td>(1.08)</td>
<td>(1.17)</td>
<td>(1.29)</td>
</tr>
<tr>
<td></td>
<td>Illiquid</td>
<td>0.06</td>
<td>0.61</td>
<td>-3.21</td>
<td>-1.41</td>
<td>-1.99</td>
<td>-2.99</td>
<td>-2.02</td>
<td>-1.96</td>
<td>-2.16</td>
<td>-2.57</td>
<td>-2.34</td>
<td>-2.66</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.01)</td>
<td>(0.03)</td>
<td>(0.34)</td>
<td>(0.49)</td>
<td>(0.54)</td>
<td>(0.62)</td>
<td>(0.67)</td>
<td>(0.72)</td>
<td>(0.75)</td>
<td>(0.80)</td>
<td>(0.84)</td>
<td>(0.89)</td>
</tr>
<tr>
<td></td>
<td>Long-Short</td>
<td>0.10</td>
<td>0.21</td>
<td>-2.26</td>
<td>-1.31</td>
<td>-2.19</td>
<td>-3.96</td>
<td>-3.40</td>
<td>-2.17</td>
<td>-2.88</td>
<td>-3.60</td>
<td>-4.24</td>
<td>-4.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.05)</td>
<td>(0.07)</td>
<td>(0.50)</td>
<td>(0.63)</td>
<td>(0.78)</td>
<td>(1.00)</td>
<td>(0.95)</td>
<td>(0.91)</td>
<td>(1.27)</td>
<td>(1.35)</td>
<td>(1.44)</td>
<td>(1.57)</td>
</tr>
</tbody>
</table>

Figure 2.7: Portfolios Formed by Turnover per Transaction. The two figures plot the simple cumulative returns reported by Table 2.9 for the zero-cost industry-neutral portfolio that is long the quintile of stocks with the highest turnover per transaction and short the quintile of stocks with the lowest turnover per transaction. The cumulative returns from the market adjusted model is on the left and from the unadjusted model is on the right. The solid line is the cumulative return of the zero-cost portfolio in percent. The dashed lines are the 95 percent confidence bands for the portfolio’s estimated return. The dotted lines are the 95 percent confidence bands for the cumulative return of an industry-neutral and size-adjusted portfolio that is long and short, respectively, 129 and 126 randomly selected stocks.
Table 2.10: Portfolios Formed by Daily Transactions. All equities listed on the Tokyo Stock Exchange are sorted by size within their respective industry groups. The smallest 30 percent of stocks and the largest 30 percent of firms within each industry group are removed to control for size effects. After removing stocks that are traded on multiple exchanges in Japan, the remaining assets are sorted by their average number of daily transactions over the period October 1, 2005 through December 31, 2005 within each of the 11 industry groups identified by Table 2.2. A zero-cost portfolio is formed by purchasing the quintile within each industry group with the highest number of transactions and selling the quintile within each industry group with the lowest number of transactions. The long and short components of the portfolio include, respectively, 129 and 126 stocks. The cumulative simple return over the d-day holding period with portfolio formation occurring on January 18, 2006 is denoted by \( R_d = \exp(\alpha_1 + \cdots + \alpha_d) - 1 \) where the alphas are the coefficients from the following regression, \( r_t = \alpha + \beta m_t + \alpha_1 \psi_1 + \cdots + \alpha_{10} \psi_{10} + \epsilon_t \), estimated over the period October 1, 2005 through February 28, 2006 where \( m_t \) is the return on the Nikkei 225 Index. The first three lines set \( I = 1 \) and are adjusted for a market factor. The following three lines set \( I = 0 \) and are unadjusted. Cumulative returns and \( \alpha \) are reported in percent. Standard errors are in parenthesis.

<table>
<thead>
<tr>
<th>Parameter Estimates</th>
<th>Portfolio</th>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>( \alpha_1 )</th>
<th>( \alpha_2 )</th>
<th>( \alpha_3 )</th>
<th>( \alpha_4 )</th>
<th>( \alpha_5 )</th>
<th>( \alpha_6 )</th>
<th>( \alpha_7 )</th>
<th>( \alpha_8 )</th>
<th>( \alpha_9 )</th>
<th>( \alpha_{10} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liquid</td>
<td>0.06</td>
<td>1.05</td>
<td>-5.70</td>
<td>-5.51</td>
<td>-5.27</td>
<td>-8.79</td>
<td>-6.67</td>
<td>-4.91</td>
<td>-6.01</td>
<td>-7.96</td>
<td>-8.19</td>
<td>-8.72</td>
<td></td>
</tr>
<tr>
<td>(0.05)</td>
<td>(0.06)</td>
<td>(0.41)</td>
<td>(0.70)</td>
<td>(0.88)</td>
<td>(1.05)</td>
<td>(0.99)</td>
<td>(1.00)</td>
<td>(1.33)</td>
<td>(1.41)</td>
<td>(1.51)</td>
<td>(1.61)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illiquid</td>
<td>0.00</td>
<td>0.43</td>
<td>-3.74</td>
<td>-1.57</td>
<td>-1.68</td>
<td>-2.49</td>
<td>-2.09</td>
<td>-2.22</td>
<td>-1.92</td>
<td>-1.90</td>
<td>-1.65</td>
<td>-1.54</td>
<td></td>
</tr>
<tr>
<td>(0.04)</td>
<td>(0.03)</td>
<td>(0.33)</td>
<td>(0.24)</td>
<td>(0.28)</td>
<td>(0.41)</td>
<td>(0.40)</td>
<td>(0.46)</td>
<td>(0.49)</td>
<td>(0.51)</td>
<td>(0.51)</td>
<td>(0.59)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long-Short</td>
<td>0.06</td>
<td>0.62</td>
<td>-1.95</td>
<td>-1.95</td>
<td>-3.59</td>
<td>-6.29</td>
<td>-4.57</td>
<td>-2.70</td>
<td>-4.10</td>
<td>-6.05</td>
<td>-6.53</td>
<td>-7.18</td>
<td></td>
</tr>
<tr>
<td>Portfolio</td>
<td>(0.06)</td>
<td>(0.07)</td>
<td>(0.75)</td>
<td>(0.94)</td>
<td>(1.13)</td>
<td>(1.08)</td>
<td>(1.11)</td>
<td>(1.42)</td>
<td>(1.51)</td>
<td>(1.61)</td>
<td>(1.73)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2.8: Portfolios Formed by Daily Transactions. The two figures plot the simple cumulative returns reported by Table 2.10 for the zero-cost industry-neutral portfolio that is long the quintile of stocks with the most daily transactions and short the quintile of stocks with the least daily transactions. The cumulative returns from the market adjusted model is on the left and from the unadjusted model is on the right. The solid line is the cumulative return of the zero-cost portfolio in percent. The dashed lines are the 95 percent confidence bands for the portfolio’s estimated return. The dotted lines are the 95 percent confidence bands for the cumulative return of an industry-neutral and size-adjusted portfolio that is long and short, respectively, 129 and 126 randomly selected stocks.
Table 2.11: Portfolios Formed by the Amihud Price Impact Illiquidity Measure. All equities listed on the Tokyo Stock Exchange are sorted by size within their respective industry groups. The smallest 30 percent of stocks and the largest 30 percent of firms within each industry group are removed to control for size effects. After removing stocks that are traded on multiple exchanges in Japan, the remaining assets are sorted by their Amihud (2002) illiquidity measure calculated over the period October 1, 2005 through December 31, 2005 within each of the 11 industry groups identified by Table 2.2. A zero-cost portfolio is formed by purchasing the quintile within each industry group with the lowest illiquidity and selling the quintile within each industry group with the highest illiquidity. The long and short components of the portfolio include, respectively, 126 and 129 stocks. The cumulative simple return over the d-day holding period with portfolio formation occurring on January 18, 2006 is denoted by \( R_\Delta \equiv \exp(\alpha_1 + \cdots + \alpha_d) - 1 \) where the alphas are the coefficients from the following regression, \( r_t = \alpha + \beta r_{m,t} + \alpha_1 \psi_1 + \cdots + \alpha_{10} \psi_{10} + \epsilon_t \), estimated over the period October 1, 2005 through February 28, 2006 where \( r_{m,t} \) is the return on the Nikkei 225 Index. The first three lines set \( I = 1 \) and are adjusted for a market factor. The following three lines set \( I = 0 \) and are unadjusted. Cumulative returns and \( \alpha \) are reported in percent. Standard errors are in parenthesis.

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>( \alpha_1 )</th>
<th>( \alpha_2 )</th>
<th>( \alpha_3 )</th>
<th>( \alpha_4 )</th>
<th>( \alpha_5 )</th>
<th>( \alpha_6 )</th>
<th>( \alpha_7 )</th>
<th>( \alpha_8 )</th>
<th>( \alpha_9 )</th>
<th>( \alpha_{10} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liquid</td>
<td>0.05</td>
<td>0.99</td>
<td>-4.93</td>
<td>-2.98</td>
<td>-4.73</td>
<td>-7.66</td>
<td>-5.99</td>
<td>-4.30</td>
<td>-5.33</td>
<td>-7.00</td>
<td>-7.25</td>
<td>-7.95</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.06)</td>
<td>(0.39)</td>
<td>(0.69)</td>
<td>(0.86)</td>
<td>(1.04)</td>
<td>(0.99)</td>
<td>(0.98)</td>
<td>(1.34)</td>
<td>(1.42)</td>
<td>(1.51)</td>
<td>(1.55)</td>
</tr>
<tr>
<td>Illiquid</td>
<td>0.00</td>
<td>0.51</td>
<td>-3.68</td>
<td>-1.57</td>
<td>-2.15</td>
<td>-3.36</td>
<td>-2.63</td>
<td>-2.62</td>
<td>-2.54</td>
<td>-2.77</td>
<td>-2.37</td>
<td>-2.60</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.04)</td>
<td>(0.34)</td>
<td>(0.27)</td>
<td>(0.38)</td>
<td>(0.53)</td>
<td>(0.50)</td>
<td>(0.52)</td>
<td>(0.55)</td>
<td>(0.56)</td>
<td>(0.64)</td>
<td>(0.67)</td>
</tr>
<tr>
<td>Long-Short</td>
<td>0.05</td>
<td>0.47</td>
<td>-1.25</td>
<td>-1.41</td>
<td>-2.58</td>
<td>-4.30</td>
<td>-3.36</td>
<td>-1.68</td>
<td>-2.79</td>
<td>-4.24</td>
<td>-4.88</td>
<td>-5.35</td>
</tr>
<tr>
<td>Portfolio</td>
<td>(0.06)</td>
<td>(0.07)</td>
<td>(0.52)</td>
<td>(0.73)</td>
<td>(0.94)</td>
<td>(1.16)</td>
<td>(1.10)</td>
<td>(1.10)</td>
<td>(1.44)</td>
<td>(1.52)</td>
<td>(1.62)</td>
<td>(1.68)</td>
</tr>
</tbody>
</table>

| Liquid    | 0.17        | -7.75       | -3.86      | -5.70      | -10.52     | -7.43      | -5.88      | -5.64      | -4.34      | -4.25      | -4.58      |
|           | (0.04)      | (0.47)      | (0.69)     | (0.87)     | (1.08)     | (1.01)     | (0.98)     | (1.33)     | (1.41)     | (1.50)     | (1.54)     |
| Illiquid  | 0.08        | -5.20       | -2.07      | -2.71      | -5.02      | -3.47      | -3.54      | -2.78      | -1.31      | -0.71      | -0.72      |
|           | (0.04)      | (0.36)      | (0.27)     | (0.38)     | (0.55)     | (0.50)     | (0.51)     | (0.54)     | (0.57)     | (0.63)     | (0.68)     |
| Long-Short| 0.10        | -2.55       | -1.79      | -2.99      | -5.50      | -3.95      | -2.34      | -2.86      | -3.03      | -3.54      | -3.86      |
| Portfolio | (0.06)      | (0.59)      | (0.74)     | (0.94)     | (1.20)     | (1.11)     | (1.10)     | (1.44)     | (1.51)     | (1.61)     | (1.67)     |

Figure 2.9: Portfolios Formed by the Amihud Price Impact Illiquidity Measure. The two figures plot the simple cumulative returns reported by Table 2.11 for the zero-cost industry-neutral portfolio that is long the quintile of stocks with the lowest Amihud (2002) illiquidity measure and short the quintile of stocks with the highest illiquidity measure. The cumulative returns from the market adjusted model is on the left and from the unadjusted model is on the right. The solid line is the cumulative return of the zero-cost portfolio in percent. The dashed lines are the 95 percent confidence bands for the portfolio’s estimated return. The dotted lines are the 95 percent confidence bands for the cumulative return of an industry-neutral and size-adjusted portfolio that is long and short, respectively, 126 and 129 randomly selected stocks.
adjusted for a market factor. The following three lines set
I through F ebruary 28, 2006 where
r portfolio’s estimated return. The dotted lines are the 95 pe rcent confidence bands for the cumulative return of an
cumulative return of the zero-cost portfolio in percent. Th e dashed lines are the 95 percent confidence bands for the
market adjusted model is on the left and from the unadjust ed model is on the right. The solid line is the
industry-neutral and size-adjusted portfolio that is long and short, respectively, 129 and 126 randomly selected stoc ks.

Table 2.12: Portfolios Formed by Y en V olume. All equities listed on the Tokyo Stock Exchange are sorted
by size within their respective industry groups. The smallest 30 percent of stocks and the largest 30 percent of firms
within each industry group are removed to control for size effects. After removing stocks that are traded on multiple
exchanges in Japan, the remaining assets are sorted by their average daily monetary volume over the period October 1,
2005 through December 31, 2005 within each of the 11 industry groups identified by Table 2.2. A zero-cost portfolio
is formed by purchasing the quintile within each industry group with the highest monetary volume and selling the
quintile within each industry group with the lowest monetary volume. The long and short components of the portfolio
include, respectively, 129 and 126 stocks. The cumulative simple return over the d-day holding period with portfolio
formation occurring on January 18, 2006 is denoted by \( R_d \equiv \exp(\alpha_1 + \cdots + \alpha_d) - 1 \) where the alphas are the coefficients
from the following regression, \( r_t = \alpha + \beta r_{m,t} + \alpha_1 \psi_1 + \cdots + \alpha_{10} \psi_{10} + \epsilon_t \), estimated over the period October 1, 2005
through February 28, 2006 where \( r_{m,t} \) is the return on the Nikkei 225 Index. The first three lines set \( I = 1 \) and are
adjusted for a market factor. The following three lines set \( I = 0 \) and are unadjusted. Cumulative returns and \( \alpha \) are reported in percent. Standard errors are in parenthesis.

<table>
<thead>
<tr>
<th>Parameter Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Portfolio</td>
</tr>
<tr>
<td>Liquid</td>
</tr>
<tr>
<td>(0.05)</td>
</tr>
<tr>
<td>Illiquid</td>
</tr>
<tr>
<td>(0.02)</td>
</tr>
<tr>
<td>Long-Short</td>
</tr>
<tr>
<td>PortFolio</td>
</tr>
</tbody>
</table>

Figure 2.10: Portfolios Formed by Y en V olume. The two figures plot the simple cumulative returns reported
by Table 2.12 for the zero-cost industry-neutral portfolio that is long the quintile of stocks with the highest monetary
trade volume and short the quintile of stocks with the lowest monetary trade volume. The cumulative returns from
the market adjusted model is on the left and from the unadjusted model is on the right. The solid line is the
cumulative return of the zero-cost portfolio in percent. The dashed lines are the 95 percent confidence bands for the
portfolio’s estimated return. The dotted lines are the 95 percent confidence bands for the cumulative return of an
industry-neutral and size-adjusted portfolio that is long and short, respectively, 129 and 126 randomly selected stocks.
I through February 28, 2006 where \( r \) from the following regression, \( R_{t} = \alpha + \beta R_{m,t} + \alpha_{1} \psi_{1} + \cdots + \alpha_{10} \psi_{10} + \epsilon_{t} \), estimated over the period October 1, 2005 through February 28, 2006 where \( R_{m} \) is the return on the Nikkei 225 Index. The first three lines set \( \psi_{i} = 1 \) and are adjusted for a market factor. The following three lines set \( \psi_{i} = 0 \) and are unadjusted. Cumulative returns and \( \alpha \) are reported in percent. Standard errors are in parenthesis.

<table>
<thead>
<tr>
<th>Parameter Estimates</th>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>( \alpha_{1} )</th>
<th>( \alpha_{2} )</th>
<th>( \alpha_{3} )</th>
<th>( \alpha_{4} )</th>
<th>( \alpha_{5} )</th>
<th>( \alpha_{6} )</th>
<th>( \alpha_{7} )</th>
<th>( \alpha_{8} )</th>
<th>( \alpha_{9} )</th>
<th>( \alpha_{10} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liquid (0.07)</td>
<td>0.10</td>
<td>1.17</td>
<td>-6.40</td>
<td>-3.93</td>
<td>-6.42</td>
<td>-10.79</td>
<td>-8.25</td>
<td>-5.95</td>
<td>-7.86</td>
<td>-9.99</td>
<td>-10.40</td>
<td>-11.14</td>
</tr>
<tr>
<td>Illiquid (0.02)</td>
<td>0.05</td>
<td>0.37</td>
<td>-3.17</td>
<td>-2.11</td>
<td>-1.31</td>
<td>-1.93</td>
<td>-1.49</td>
<td>-2.13</td>
<td>-1.96</td>
<td>-1.80</td>
<td>-1.51</td>
<td>-1.58</td>
</tr>
<tr>
<td>Long-Short Portfolio (0.07)</td>
<td>0.05</td>
<td>0.80</td>
<td>-3.23</td>
<td>-2.72</td>
<td>-5.11</td>
<td>-8.86</td>
<td>-6.76</td>
<td>-3.82</td>
<td>-5.90</td>
<td>-8.19</td>
<td>-8.89</td>
<td>-9.56</td>
</tr>
</tbody>
</table>

| Liquid (0.07) | 0.24 | -9.69 | -4.95 | -7.53 | -14.05 | -9.90 | -7.77 | -8.20 | -6.92 | -6.96 | -7.33 |
| Illiquid (0.02) | 0.10 | -4.27 | -1.56 | -1.72 | -3.16 | -2.11 | -2.80 | -2.14 | -0.76 | -0.32 | -0.23 |
| Long-Short Portfolio (0.07) | 0.14 | -5.41 | -3.39 | -5.81 | -10.89 | -7.79 | -4.97 | -6.06 | -6.17 | -6.64 | -7.11 |

Table 2.13: Portfolios Formed by Turnover and Yen Volume. All equities listed on the Tokyo Stock Exchange are sorted by size within their respective industry groups. The smallest 30 percent of firms within each industry group are removed to control for size effects. After removing stocks that are traded on multiple exchanges in Japan, the remaining assets are sorted separately by their average daily turnover and by their average daily monetary volume over the period October 1, 2005 through December 31, 2005 within each of the 11 industry groups identified by Table 2.2. A zero-cost portfolio is formed by purchasing the assets in each industry group within the highest turnover quintile and highest monetary volume quintile and selling assets within the lowest turnover quintile and lowest monetary volume quintile. The long and short components of the portfolio include, respectively, 101 and 93 stocks. The cumulative simple return over the \( d \)-day holding period with portfolio formation occurring on January 18, 2006 is denoted by \( R_{d} = \exp(\alpha_{1} + \cdots + \alpha_{d}) - 1 \) where the alphas are the coefficients from the following regression, \( R_{t} = \alpha + \beta R_{m,t} + \alpha_{1} \psi_{1} + \cdots + \alpha_{10} \psi_{10} + \epsilon_{t} \), estimated over the period October 1, 2005 through February 28, 2006 where \( R_{m} \) is the return on the Nikkei 225 Index. The first three lines set \( \psi_{i} = 1 \) and are adjusted for a market factor. The following three lines set \( \psi_{i} = 0 \) and are unadjusted. Cumulative returns and \( \alpha \) are reported in percent. Standard errors are in parenthesis.

Figure 2.11: Portfolios Formed by Turnover and Yen Volume. The two figures plot the simple cumulative returns reported by Table 2.13 for the zero-cost industry-neutral portfolio that is long the stocks with high turnover and high monetary trade volume and short stocks with low turnover and low monetary trade volume. The cumulative returns from the market adjusted model is on the left and from the unadjusted model is on the right. The solid line is the cumulative return of the zero-cost portfolio in percent. The dashed lines are the 95 percent confidence bands for the cumulative return of an industry-neutral and size-adjusted portfolio that is long and short, respectively, 101 and 93 randomly selected stocks.
Table 2.14: Portfolios Formed by Turnover and the Amihud (2002) Illiquidity Measure. All equities listed on the Tokyo Stock Exchange are sorted by size within their respective industry groups. The smallest 30 percent of stocks and the largest 30 percent of firms within each industry group are removed to control for size effects. After removing stocks that are traded on multiple exchanges in Japan, the remaining assets are sorted separately by their average daily turnover and by their Amihud (2002) illiquidity measure calculated over the period October 1, 2005 through December 31, 2005 within each of the 11 industry groups identified by Table 2.2. A zero-cost portfolio is formed by purchasing the assets in each industry group within the highest turnover quintile and lowest illiquidity quintile and selling assets within the lowest turnover quintile and highest illiquidity quintile. The long and short components of the portfolio include, respectively, 80 and 82 stocks. The cumulative simple return over the \( d \)-day holding period with portfolio formation occurring on January 18, 2006 is denoted by \( R_d = \exp(\alpha_1 + \cdots + \alpha_d) - 1 \) where the alphas are the coefficients from the following regression, \( r_t = \alpha + \beta r_{m,t} + \alpha_1 \psi_1 + \cdots + \alpha_{10} \psi_{10} + \epsilon_t \), estimated over the period October 1, 2005 through February 28, 2006 where \( r_{m,t} \) is the return on the Nikkei 225 Index. The first three lines set \( I = 1 \) and are adjusted for a market factor. The following three lines set \( I = 0 \) and are unadjusted. Cumulative returns and \( \alpha \) are reported in percent. Standard errors are in parenthesis.

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>( \alpha_1 )</th>
<th>( \alpha_2 )</th>
<th>( \alpha_3 )</th>
<th>( \alpha_4 )</th>
<th>( \alpha_5 )</th>
<th>( \alpha_6 )</th>
<th>( \alpha_7 )</th>
<th>( \alpha_8 )</th>
<th>( \alpha_9 )</th>
<th>( \alpha_{10} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liquid</td>
<td>0.10</td>
<td>1.22</td>
<td>-6.28</td>
<td>-4.16</td>
<td>-6.67</td>
<td>-10.69</td>
<td>-8.30</td>
<td>-6.02</td>
<td>-7.66</td>
<td>-10.01</td>
<td>-10.50</td>
<td>-11.46</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.52)</td>
<td>(1.04)</td>
<td>(1.29)</td>
<td>(1.50)</td>
<td>(1.46)</td>
<td>(1.46)</td>
<td>(2.03)</td>
<td>(2.13)</td>
<td>(2.27)</td>
<td>(2.31)</td>
</tr>
<tr>
<td>Illiquid</td>
<td>0.05</td>
<td>0.36</td>
<td>-3.29</td>
<td>-1.44</td>
<td>-1.60</td>
<td>-2.35</td>
<td>-2.01</td>
<td>-2.87</td>
<td>-2.73</td>
<td>-2.43</td>
<td>-2.12</td>
<td>-2.36</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.03)</td>
<td>(0.38)</td>
<td>(0.29)</td>
<td>(0.33)</td>
<td>(0.49)</td>
<td>(0.51)</td>
<td>(0.57)</td>
<td>(0.55)</td>
<td>(0.53)</td>
<td>(0.58)</td>
<td>(0.63)</td>
</tr>
<tr>
<td>Long-Short Portfolio</td>
<td>0.05</td>
<td>0.86</td>
<td>-2.99</td>
<td>-2.72</td>
<td>-5.08</td>
<td>-8.33</td>
<td>-6.29</td>
<td>-3.14</td>
<td>-4.93</td>
<td>-7.59</td>
<td>-8.38</td>
<td>-9.09</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.08)</td>
<td>(0.64)</td>
<td>(1.07)</td>
<td>(1.32)</td>
<td>(1.57)</td>
<td>(1.53)</td>
<td>(1.55)</td>
<td>(2.08)</td>
<td>(2.17)</td>
<td>(2.31)</td>
<td>(2.37)</td>
</tr>
</tbody>
</table>

Figure 2.12: Portfolios Formed by Turnover and the Amihud (2002) Illiquidity Measure. The two figures plot the simple cumulative returns reported by Table 2.13 for the zero-cost industry-neutral portfolio that is long the stocks with high turnover and large number of daily transactions and short stocks with low turnover and low number of daily transactions. The cumulative returns from the market adjusted model is on the left and from the unadjusted model is on the right. The solid line is the cumulative return of the zero-cost portfolio in percent. The dashed lines are the 95 percent confidence bands for the portfolio’s estimated return. The dotted lines are the 95 percent confidence bands for the cumulative return of an industry-neutral and size-adjusted portfolio that is long and short, respectively, 80 and 82 randomly selected stocks.

![Market Adjusted](image)

![Not Market Adjusted](image)
Table 2.15: Estimates of Bid-Ask Spreads. Bid-ask spreads are estimated for each stock listed on the Tokyo Stock Exchange using the Gibbs Sampler Estimator suggested by Hasbrouck (2006) applied to transactions level data. Three separate sample windows are considered. The first window is the ten trading day period prior to the market closure. The second window is the ten trading day period including and subsequent to the market closure. The final window is the next ten trading days. Stocks with less than 100 observations per window are excluded. Stocks whose estimates fall in the upper and lower 1 percent tails within any of the three windows are also excluded. Panel A presents the summary statistics for stocks traded on TSE. Panel B reports the summary statistics for stocks traded only on TSE. Panel C reports the summary statistics for stocks traded on TSE and at least one additional Japanese exchange. The number in parentheses is the count of stocks meeting the inclusion requirements within the respective category. The lower and upper breakpoints for the first and tenth decile are provided by p(10) and p(90), forming the 80 percent confidence region. All numbers are reported in basis points.

<table>
<thead>
<tr>
<th>Period</th>
<th>All stocks traded on TSE (2078)</th>
<th>Stocks Traded Only on TSE (1501)</th>
<th>Cross-Listed Stocks Traded on TSE (577)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Period</td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>12/30/2005 – 1/17/2006</td>
<td>23.4</td>
<td>20.7</td>
<td>14.2</td>
</tr>
<tr>
<td>1/18/2006 – 1/31/2006</td>
<td>26.2</td>
<td>22.7</td>
<td>15.9</td>
</tr>
<tr>
<td>2/1/2006 – 2/14/2006</td>
<td>23.6</td>
<td>20.3</td>
<td>15.8</td>
</tr>
</tbody>
</table>

Table 2.16: Changes in Spreads. This table reports the systematic changes in bid-ask spreads across periods surrounding the market closure on January 18, 2006. The first period is for the ten-day window of 12/31/2005 through 1/17/2006. The second period is for the ten-day window of 1/18/2006 through 1/31/2006. The third period is for the ten-day window of 2/1/2006 through 2/14/2006.

<table>
<thead>
<tr>
<th>Period</th>
<th>All (2078)</th>
<th>TSE (1501)</th>
<th>TSE+ (577)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2nd – 1st</td>
<td>2.80</td>
<td>3.03</td>
<td>2.21</td>
</tr>
<tr>
<td></td>
<td>(0.21)</td>
<td>(0.26)</td>
<td>(0.32)</td>
</tr>
<tr>
<td>3rd – 2nd</td>
<td>-2.55</td>
<td>-2.46</td>
<td>-2.80</td>
</tr>
<tr>
<td></td>
<td>(0.23)</td>
<td>(0.29)</td>
<td>(0.37)</td>
</tr>
<tr>
<td>3rd – 1st</td>
<td>0.25</td>
<td>0.57</td>
<td>-0.57</td>
</tr>
<tr>
<td></td>
<td>(0.23)</td>
<td>(0.30)</td>
<td>(0.32)</td>
</tr>
</tbody>
</table>

94
Chapter 3

Penalized Method of Moments and Empirical Likelihood Estimation: Negative Weights and Shrinkage

3.1 Introduction

Economists frequently wish to estimate a vector of unknown parameters using a set of Euler equations with expectation zero. Two-step Generalized Method of Moments (GMM) is a widely used tool employed in estimating unknown parameters. When the number of equations equals the number of parameters the system is just identified and the econometrician is able to select a parameter vector that sets the estimation equations equal to zero. In a just identified system, the model has not been tested – the econometrician has only described the data. The economic model may be tested by providing overidentifying restrictions through additional estimation equations. One would hope that the additional information provided by these equations would lead to better estimator properties. It is well known that increasing the number of overidentifying restrictions in small samples adversely affects estimator bias and leads to mis-sized hypothesis tests.

In order to improve the small-sample properties of GMM estimation, a number of researchers have proposed and investigated a variety of one-step alternatives to GMM. Three approaches that have received special attention are the Empirical Likelihood (EL) estimator of Owen (1988), Qin and Lawless (1994), and Imbens (1997), the Exponential Tilting (ET)
estimator of Kitamura and Stutzer (1997) and Imbens, Johnson, and Spady (1998), and the Continuous Upd ating GMM (CUE) estimator of Hansen, Heaton, and Yaron (1996). Newey and Smith (2004) show that the three estimators are members of the Cressie and Read (1984) (CR) family of discrepancies applied to the Minimum Discrepancy (MD) estimation formulated by Corcoran (1998). The essence of MD estimation is to allow the sample weights applied to the estimation of the expected moment vector to deviate from GMM’s fixed weights and to jointly select the weights and parameter vector that achieve minimum empirical discrepancy (from fixed weights) subject to the constraint that the sample average moment vector equals zero.

This paper suggests an alternate approach to MD-style estimation. I propose an estimator that combines the flexibility offered by GMM, which allows the moment vector to be non-zero, with that of EL, which allows the weighting vector to be non-fixed. The sample average moment vector is allowed to deviate from zero, but the deviation is costly through a GMM-type quadratic penalty function. The weighting vector is allowed to deviate from $n^{-1}$, but the deviation is costly through EL’s Kullback-Leibler Information Criterion (KLIC) penalty function. The resulting estimation procedure, which I name Penalized Method of Moments (PMM), has an objective function that resembles a combination of the GMM and EL objective functions. By including a parameter, $\delta \in (0, 1)$, PMM becomes a continuum of estimators with behavior approaching GMM when $\delta$ approaches zero and with behavior approaching EL when $\delta$ approaches one.

EL’s objective function is undefined for non-positive sample weights. This property results in the possibility that the estimator is undefined at or near the population parameter value, a possibility that increases in probability as the sample size decreases, or the number of moments increase, or with model misspecification. The PMM estimator, which is also undefined for non-positive sample weights, allows the econometrician to increase the support of the parameter space by shrinking the optimal weights towards GMM’s fixed weights through $\delta$. This feature provides the econometrician with an additional degree of flexibility under conditions when a restricted parameter space is likely to be problematic.

I perform a series of Monte Carlo experiments on a Hall and Horowitz (1996) style model.
as modified and investigated by Schennach (2006). I compare certain properties of PMM estimation, such as the size of hypothesis tests and estimator volatility, to those of EL for the specific estimator obtained by setting $\delta = 0.5$. The simulations are performed on small samples (25, 50, and 100 observations) with a large number of moment conditions to investigate the properties of PMM when the restricted parameter space is likely to be problematic.

I report that under a small sample with 1 overidentifying restriction, PMM and EL provide estimates with comparable variability and similarly mis-sized hypothesis tests. When the sample size is larger and the number of moment conditions are small, EL’s higher order properties are evidenced by its estimates being less variable than those provided by PMM. When the number of moment conditions is large relative to the sample size, however, the effect of the restricted parameter space is evident. In these experiments, PMM’s estimates are significantly less volatile and its hypothesis tests are less mis-sized than those of EL. The experiments suggest that, in certain circumstances, the restricted parameter space affects the properties of the estimates and when this is the case, PMM provides a mechanism to reduce the negative impact by increasing the support of the parameter space.

The increased support of the parameter space provided by PMM also offers a practical advantage for operationalizing the estimator. In order to begin the numerical optimization, the econometrician must supply a starting value. Unfortunately, it is possible that EL and PMM are undefined at the supplied starting parameter vector because of the restricted parameter space. When this occurs, the econometrician must search for a starting value in which the estimators are defined, a task that becomes increasingly difficult as the number of parameters increases. The larger support provided by PMM increases the likelihood and decreases the difficulty in finding an appropriate starting value.

The remainder of the paper is organized as follows. The next section provides the notation used throughout the paper. Section 3 summarizes GMM and EL. PMM is introduced in Section 4. I describe the CID phenomenon in Section 5. Section 6 briefly describes the role of the free parameter for the PMM family of estimators. I describe how to implement PMM in Section 7. The results of the Monte Carlo study is reported in Section 8. Section 9 concludes. The Appendix provides all proofs.
3.2 Definitions and Notation

This section provides all the notation and definitions used throughout the paper. Let \( x_i \) be a sequence of i.i.d. random vectors taking values in \( \mathcal{X} \subset \mathbb{R}^p \). Let \( \theta \) denote the parameter vector of interest belonging in the space \( \Theta \), a compact subset of \( \mathbb{R}^k \). Our goal is to estimate the population parameter vector, denoted \( \theta_0 \), using \( g(x_i, \theta) \), a vector of estimating equations taking values in \( \mathbb{R}^m \) with \( m \geq k \) and satisfying

\[
E[g(x_i, \theta_0)] = 0 \quad \quad E[g(x_i, \theta_0)g(x_i, \theta_0)'] = \Sigma. \quad (3.1)
\]

Let \( n \) denote the sample size and let all summations be over \( 1, \ldots, n \). Let \( \omega^\theta, \omega^m, \) and \( \omega^v \) denote vectors of weights on the \((n - 1)\) dimensional unit simplex given by \( \Psi = \{ \omega = (\omega_1, \ldots, \omega_n) | \omega_i \geq 0, \sum_i \omega_i = 1 \} \). Let \( g_i(\theta) \) represent \( g(x_i, \theta) \) and define \( m_i(\theta) \equiv \partial g_i(\theta)/\partial \theta' \), \( v_i(\theta) \equiv g_i(\theta)g_i(\theta)' \), and

\[
G_n(\theta) = \frac{1}{n} \sum_i g_i(\theta) \quad \quad \mathcal{G}_n(\theta) \equiv \mathcal{G}_n(\omega, \theta) = \sum_i \omega_i^\theta g_i(\theta) \\
M_n(\theta) = \frac{1}{n} \sum_i m_i(\theta) \quad \quad \mathcal{M}_n(\theta) \equiv \mathcal{M}_n(\omega, \theta) = \sum_i \omega_i^m m_i(\theta) \\
V_n(\theta) = \frac{1}{n} \sum_i v_i(\theta) \quad \quad \mathcal{V}_n(\theta) \equiv \mathcal{V}_n(\omega, \theta) = \sum_i \omega_i^v v_i(\theta) - \mathcal{G}_n(\theta)\mathcal{G}_n(\theta)'
\]

In addition, \( G, \mathcal{G}, M, \mathcal{M}, V, \) and \( \mathcal{V} \) refer to limiting value as \( n \to \infty \) of the respective functions evaluated at the true parameter value \( \theta_0 \). Let \( W_n \) refer to an \( m \times m \) full rank symmetric positive definite weighting matrix with limiting value \( W \) as \( n \to \infty \). Let \( P_{M}^\perp = I_m - \Sigma^{-\frac{1}{2}} M (M'\Sigma^{-1} M)^{-1} M'\Sigma^{-\frac{1}{2}} \) be the projection matrix orthogonal to the space spanned by the asymptotic normalized Jacobian, where \( \Sigma^{-\frac{1}{2}} \) is the Cholesky decomposition of \( \Sigma^{-1} \). Define \( \Omega \equiv M'\Sigma^{-1} M, \tilde{\Omega} \equiv M'\tilde{W} M, \Upsilon \equiv \Omega^{-1} M'\Sigma^{-1} \), and \( \check{\Upsilon} \equiv \check{\Omega}^{-1} M'\check{W} \), where \( \check{W} \) is a first-round weighting matrix. Finally, the following four definitions are related to the
higher-order bias and are referenced by all the estimators discussed in this paper:

\[ B_I = n^{-1} \Upsilon (E [m_i(\theta_0)\Upsilon g_i(\theta_0)] - a/2) \]  
(3.2)

\[ B_M = -n^{-1} \Omega^{-1} \E \left[ m_i(\theta_0)^\prime \Sigma^{-\frac{1}{2}} P^\perp \Sigma^{-\frac{1}{2}} g_i(\theta_0) \right] \]  
(3.3)

\[ B_\Sigma = n^{-1} \Upsilon \E \left[ v_i(\theta_0)^\prime \Sigma^{-\frac{1}{2}} P^\perp \Sigma^{-\frac{1}{2}} g_i(\theta_0) \right] \]  
(3.4)

\[ B_W = -n^{-1} \Upsilon \E \left[ \sum_{j=1}^{k} \frac{\partial V_n(\theta_0)}{\partial \theta_j} \right] (\tilde{\Upsilon} - \Upsilon)^\prime e_j, \]  
(3.5)

where \( e_j \) is the \( j \)th unit vector, \( a \) is an \( m \times 1 \) vector with

\[ a_j \equiv \text{tr} \left( \Omega \E \left[ \frac{\partial^2 g_{ij}(\theta_0)}{\partial \theta \partial \theta^\prime} \right] \right) \]  
(3.6)

and \( g_{ij}(\theta_0) \) denotes the \( j \)th element of \( g_i(\theta_0) \).

Finally, let \( f(n) = O(g(n)) \) denote an asymptotic upper bound

\[ f(n) = O(g(n)) \iff \limsup_{n \to \infty} \left| \frac{f(n)}{g(n)} \right| < \infty \]  
(3.7)

and \( f(n) = o(g(n)) \) denote asymptotic negligibility

\[ f(n) = o(g(n)) \iff \lim_{n \to \infty} \frac{f(n)}{g(n)} = 0. \]  
(3.8)

3.3 The Ingredients: GMM and EL

The family of estimators proposed in this paper merges the objective functions of GMM and EL. Before providing the details of the new estimator, I provide a summary of its two ingredients for comparison purposes.

3.3.1 Generalized Method of Moments

Moment estimation is based on a set of estimating equations with expectation zero. The objective is to select a value \( \hat{\theta} \) that is consistent with the information contained in the mo-
ment conditions. GMM uses fixed weights to estimate the expected value of the estimating equations. Because the sample weights are fixed to \( n^{-1} \), when there are more moments than parameters it is not possible to jointly satisfy the \( m \) moment conditions. GMM sets \( k \) dimensions of the moment vector to zero and measures the distance from zero of the remaining \( m - k \) dimensions with a quadratic penalty function. Optimality is defined by the parameter value associated with the lowest penalty. The GMM estimate of \( \hat{\theta} \) is the solution to the following optimization problem:

**Definition 1 (GMM estimator)**

\[
\hat{\theta} = \arg \min_{\theta \in \Theta} Q(\theta) = nG_n(\theta)'W_nG_n(\theta),
\]

(3.9)

where \( W_n \) is a symmetric positive definite \( m \times m \) weighting matrix. The first-order conditions for GMM are

\[
M_n(\hat{\theta})'W_nG_n(\hat{\theta}) = 0,
\]

(3.10)

which is a \( k \times 1 \) vector that represents the identifying space for the \( k \) parameters. GMM is estimated efficiently by setting \( W_n \overset{p}{\rightarrow} \Sigma^{-1} \), which is typically accomplished by estimating equation (3.9) twice. The first step uses a user supplied weighting matrix \( \tilde{W} \), such as the identity matrix, and provides the first round consistent estimate \( \tilde{\theta} \). The second step sets \( W_n = V_n^{-1}(\tilde{\theta}) \), which is a consistent estimate of \( \Sigma^{-1} \). Under standard regularity conditions, the estimator \( \hat{\theta} \) is consistent and asymptotically normally distributed. The asymptotic distribution of the estimator is

\[
\sqrt{n}(\hat{\theta} - \theta_0) \overset{d}{\rightarrow} N \left( 0, (M'WM)^{-1}M'W\Sigma W M(M'WM)^{-1} \right).
\]

(3.11)

If an efficient weighting matrix is used, then the asymptotic distribution reduces to

\[
\sqrt{n}(\hat{\theta} - \theta_0) \overset{d}{\rightarrow} N \left( 0, (M'\Sigma^{-1}M)^{-1} \right)
\]

(3.12)

and the overidentifying restrictions may be tested using the \( J \)-test statistic proposed by
Hansen (1982),

\[ Q(\hat{\theta})_{\text{GMM}} \overset{d}{\rightarrow} \chi^2_{m-k}. \]  

(3.13)

Newey and Smith (2004) show that GMM’s higher-order \( O(n^{-1}) \) bias is

\[
\text{Bias}(\hat{\theta})_{\text{GMM}} = B_I + B_M + B_\Sigma + B_W.
\]  

(3.14)

The first component, \( B_I \), is the bias for an estimator with the optimal (non-random) transformation \( M'\Sigma^{-1} \) applied to the moment conditions \( G_n(\theta_0) \). The second component and third components, \( B_M \) and \( B_\Sigma \), are the bias due to the estimation of the Jacobian and second moment matrices. The final term, \( B_W \), is the bias due to the choice of the first-round weighting matrix.

3.3.2 Empirical Likelihood

EL takes a different approach to estimation. EL calculates the sample average moment vector with weights that are allowed to deviate from GMM’s fixed weights and imposes the constraint that the sample average moment vector equals zero. Because sample weights are not fixed, it is possible to select a vector of non-negative probability weights under which the expected value of the estimating equations is equal to zero so long as zero is contained in the convex hull of the points \( g(x_1, \theta), \ldots, g(x_n, \theta) \). EL is defined over the set of weights and parameter values such that this condition is satisfied (i.e. \( G_n(\theta) = 0 \) and \( \omega > 0 \)). When the system is overidentified, EL’s moment vector constraint may not be satisfied under fixed weights. EL measures the distance from \( n^{-1} \) with a penalty function defined by the Kullback-Leibler Information Criterion (KLIC) and optimality is defined by the parameter and weighting vectors associated with the lowest penalty. When the system is just identified, it is possible to set the expected moment condition to zero under fixed weights, which results in the same solution provided by GMM estimation. EL is defined below.
Definition 2 (EL Estimator)

\[ \hat{\theta} = \arg \min_{\theta \in \Theta, \omega \in \Psi} Q(\omega) = -2 \sum_i \ln n \omega_i \quad \text{subject to} \quad G_n(\theta) = 0. \]  

(3.15)

EL minimizes the empirical discrepancy as measured by the KLIC. An alternative view is that EL maximizes the likelihood of the multinomial distribution for the data subject to the constraint that the orthogonality conditions are binding. Unlike GMM, estimating with EL requires no iteration to achieve efficiency. The probability weights that satisfy the first-order conditions are the solution to the following system of equations

\[ \omega_i = \frac{1}{n} \left( \frac{1}{1 + g_i(\theta) V_n^{-1}(\theta) G_n(\theta)} \right). \]  

(3.16)

The weights cannot be solved explicitly because \( V_n(\theta) \) is a function of the probability weights. Denote the implicit solution to (3.16) by \( \omega_i(\theta) \). These optimal weights are conditional on \( \theta \) and may be negative. In order to satisfy the restriction that \( \omega \in \Psi \), the estimation procedure restricts the parameter space to the region in which the optimal weights are positive. If this region is the null space, then the EL estimator is undefined for the combined model and data set. When this is the case, a procedure such as the shrinkage factor approach of Bonnal and Renault (2004) may be implemented. Their approach, which is asymptotically equivalent to EL, shrinks the optimal weights towards fixed weights such that all negative weights are removed.

Newey and Smith (2004) show that the EL estimator \( \hat{\theta} \) is selected as the solution to the following first-order conditions

\[ \mathcal{M}_n(\hat{\theta}) V_n^{-1}(\theta) G_n(\hat{\theta}) = 0. \]  

(3.17)

Equation (3.17) is a \( k \times 1 \) vector that represents the identifying space for the \( k \) parameters and is a function of the probability weights. Hence equations (3.16) and (3.17) must be solved jointly for the \( n \) weights and the \( k \) parameters. Note the similar structure between the first-order conditions provided by equations (3.17) and (3.10). In both sets of first-order
conditions a linear combination of the sample average moment vector is set to zero. The linear combination is comprised of estimates of the Jacobian and the second moment matrix. The difference is that for EL, the expected Jacobian is calculated using the probability weights that minimize its objective function with the weighting matrix calculated simultaneously. GMM, on the other hand, uses fixed weights to calculate the average score and the weighting matrix is supplied \textit{a priori}.

Jointly solving for \( n \) weights and \( k \) parameters is a computationally formidable task. It is possible to reduce the task from \( n + k \) to \( m + k \) dimensions. The system of equations that define the optimal weights may be rewritten in terms of an unknown \( m \times 1 \) vector \( \vartheta \):

\[
\omega_i = \frac{1}{n} \left( \frac{1}{1 + g_i(\theta)^\prime \vartheta} \right).
\]  

Then, equation (3.18) and the moment constraint \( G_n(\theta) = 0 \) provide \( m \) equations and \( m \) unknowns. The vector \( \vartheta(\theta) \) is solved as in implicit function of the parameter vector and equation (3.18) provides the optimal weights as an implicit function of the parameter vector.

The following statistic, provided by Qin and Lawless (1994) and Newey and Smith (2004), may be used to test for overidentifying restrictions:

\[
nG_n(\hat{\theta})'V_n^{-\frac{1}{2}}(\hat{\theta})P_{M}^{-1}(-)V_n^{-\frac{1}{2}}(\hat{\theta})G_n(\hat{\theta}) \xrightarrow{d} \chi^2_{m-k},
\]  

where \((-\)) represents a generalized inverse. Alternatively, Qin and Lawless (1994) show that the likelihood ratio test statistic may be used:

\[
Q(\omega(\hat{\theta})) \xrightarrow{d} \chi^2_{m-k}.
\]

Newey and Smith (2004) show that EL’s higher-order \( \mathcal{O}(n^{-1}) \) bias is

\[
\text{Bias}(\hat{\theta}) = B_L.
\]

If the higher-order bias of EL is compared against that of GMM, which is provided by
equation (3.14), we see that three of the four components disappear. EL provides efficient estimates of the Jacobian and second moment matrices, in that its estimates do not contribute to the higher order bias. As a one-step estimator, EL has no bias component from the inclusion of a preliminary weighting matrix. We must be careful about our interpretation of this comparison of the two estimators. Although EL certainly has fewer higher-order bias contributors, EL may not have the smaller higher-order bias for a given estimation problem. The four bias-terms are not guaranteed to have the same sign and it is possible that the three missing bias terms help to cancel out some of the $B_I$ bias.

3.4 Penalized Method of Moments

In the limit, the equal-weight sample average moment vector converges to its expected value, zero. But the sample average moment vector is a random vector, and in small samples adding flexibility by allowing some deviation from the asymptotically optimal values may provide desirable estimation properties. For instance, one strategy of robust estimation is to alter the weights applied to individual observations by removing or downweighting ‘outliers’. As described in the previous section, GMM restricts the weights to $n^{-1}$ and sets $k$ sample moments to zero; $m - k$ sample moments are not restricted and may deviate from zero. EL allows the weights to deviate from $n^{-1}$ and restricts the $m$ sample moments to zero. I propose an estimator that allows both the weights and the moment vector to deviate from $n^{-1}$ and zero respectively. Deviation of the weighted-average moment vector is costly through a quadratic penalty function and deviation of the sample weights is costly through a KLIC penalty function. The Penalized Method of Moments (PMM) estimator is defined below:

Definition 4 (PMM Estimator)

$$
\hat{\theta} = \arg \min_{\theta \in \Theta} \min_{\omega \in \Psi} Q(\omega, \theta) = \frac{1}{\delta(1 - \delta)} \left[ \delta n G_n(\theta) W_n G_n(\theta) - 2(1 - \delta) \sum_i \ln(n \omega_i) \right],
$$

(3.22)

where $\delta \in (0, 1)$ allows for the relative importance of the quadratic penalty versus the KLIC
penalty to be adjusted. The division by $\delta(1 - \delta)$ ensures proper scaling so that $Q(\hat{\omega}, \hat{\theta})$ may serve as a statistic for the test of overidentifying restrictions.

In order to solve the optimization problem, I follow the approach of Qin and Lawless (1994) and use Lagrange multipliers. Let

$$L = \frac{1}{\delta(1 - \delta)} \left[ \delta n G_n(\theta)' W_n G_n(\theta) - (1 - \delta)2 \sum_i \ln(n\omega_i) \right] - \lambda \left( \sum_i \omega_i - 1 \right), \quad (3.23)$$

where $\lambda$ is the Lagrange multiplier. Differentiating with respect to $\omega_i$, I obtain the first-order conditions

$$\frac{\partial L}{\partial \omega_i} = \frac{2}{\delta(1 - \delta)} \left[ \delta n g_i(\theta)' W_n G_n(\theta) - \frac{1 - \delta}{\omega_i} \right] - \lambda = 0. \quad (3.24)$$

Multiplying by $\omega_i$ and summing over $i$ provides

$$\sum_i \omega_i \frac{\partial L}{\partial \omega_i} = \frac{2}{\delta(1 - \delta)} \left[ \delta n G_n(\theta)' W_n G_n(\theta) - n(1 - \delta) \right] - \lambda = 0. \quad (3.25)$$

Solving for $\lambda$ yields

$$\lambda = \frac{2}{\delta(1 - \delta)} \left[ n \delta G_n(\theta)' W_n G_n(\theta) - n(1 - \delta) \right]. \quad (3.26)$$

Substituting $\lambda$ from equation (3.26) into (3.24) and rearranging terms provides the following system of $n$ equations that identify the $n$ probability weights:

$$\omega_i = \frac{1}{n} \left( \frac{1}{1 + \frac{\delta}{1 - \delta} (g_i(\theta) - G_n(\theta))' W_n G_n(\theta)} \right). \quad (3.27)$$

For a number of calculations remaining in the paper, it will be convenient to rewrite the weights as

$$\omega_i = \frac{1}{n} - \frac{1}{n} \left( \frac{\frac{\delta}{1 - \delta} (g_i(\theta) - G_n(\theta))' W_n G_n(\theta)}{1 + \frac{\delta}{1 - \delta} (g_i(\theta) - G_n(\theta))' W_n G_n(\theta)} \right). \quad (3.28)$$

Equations (3.27) and (3.28) are not closed form solutions for the weights because $G_n(\theta)$ is a function of the weight vector. However, the $n$ equations imply a solution to the $n$ unknowns conditional on $\theta$, and the implied weights, $\omega(\theta)$, are a function of the parameter vector. In the Appendix, I verify through application of the implicit function theorem that an implicit
function $\omega(\theta)$ exists asymptotically. For PMM, all three components, $G_n(\theta)$, $M_n(\theta)$, and $V_n(\theta)$ use the same vector of weights.

Because the optimal PMM weights are an implicit function of the parameter vector, the optimally weighted objective function may be written as a function of only the parameter vector:

$$Q_{PMM}(\hat{\omega}, \theta) = Q(\omega(\theta), \theta) = Q(\theta).$$

So far, only the optimal weights have been defined. The optimal parameter vector is obtained by differentiating (3.23) with respect to $\theta$. The first-order conditions imply

$$M_n(\hat{\theta})' W_n G_n(\hat{\theta}) = 0. \quad (3.29)$$

Equation (3.29) is a system of $k$ equations with $k$ unknowns. The first-order conditions are a function of the optimal weights, but the optimal weights are implicitly a function of the parameter vector. Hence, equations (3.27) and (3.29) are solved simultaneously for the $n$ optimal weights and the $k$ optimal parameters. The first-order conditions for PMM are in the familiar form in which a linear combination of an estimate of the orthogonality condition is set to zero. The linear combination is comprised of an estimate of the Jacobian term and a weighting matrix. In the previous section, I noted that (i) GMM estimates the Jacobian using sample averages and uses a weighting matrix provided by the econometrician and (ii) EL uses optimally weighted estimates of both the Jacobian term and the second moment weighting matrix. Newey and Smith (2004) show that for the continuum of estimators in the Generalized Empirical Likelihood (GEL) framework, the moment conditions $G_n(\hat{\theta})$ are always computed by sample averages while the Jacobian and second moment matrix may be calculated using different optimally weighted averages that depend on the estimator. The PMM first-order conditions may appear to be different because a weighted average of the moment conditions is used instead of the sample average. I will show that this is not exactly the case.

Define $S_n(\theta)$ as

$$S_n(\theta) \equiv (\delta V_n(\theta) + (1 - \delta)W_n^{-1})^{-1}. \quad (3.30)$$
The matrix $S_n(\theta)$ is a continuously updated weighting matrix. Its inverse is a convex combination of the efficiently estimated second moment matrix and the inverted weighting matrix provided by the econometrician. Efficient GMM sets the inverted weighting matrix equal to a consistent estimate of $\Sigma$. If the weighting matrix satisfies this condition, then the $S_n^{-1}(\hat{\theta})$ is an estimate of $\Sigma$.

**Lemma 3.4.1** Suppose $\delta \in (0,1)$, $W_n$ is a full rank positive definite symmetric $m \times m$ matrix, and $\omega(\theta)$ is the vector of weights implied by the first order conditions of equation (3.23) and defined by (3.28). Then

$$G_n(\theta) = (1 - \delta)W_n^{-1}S_n(\theta)G_n(\theta)$$  \hspace{1cm} (3.31)

Lemma 3.4.1 relates the optimally weighted to the equally weighted sample average moment vector for any $\theta$. The two estimates are explicitly related and $\delta$ has an important role (that will be further discussed) in the relationship. Lemma 3.4.1 may be used to provide a second view of the first-order conditions that simultaneously define $\hat{\theta}$ and the optimal weights. Substituting equation (3.31) into equation (3.29) provides an alternative, but equivalent set of first-order conditions:

$$M_n(\hat{\theta})'S_n(\hat{\theta})G_n(\hat{\theta}) = 0.$$  \hspace{1cm} (3.32)

Equation (3.29) shows that PMM’s first-order conditions have the same structure as all the estimators in the CR family – a linear combination of the equal-weight sample average moment vector is set to zero. The first-order conditions depend on the weights through $M_n(\hat{\theta})$ and $S_n(\hat{\theta})$ and the optimal parameter vector and optimal weights are estimated simultaneously. In contrast to the CR family of one-step estimators, the PMM first-order conditions demonstrate the two-step approach through $S_n(\theta)$’s dependance on $W_n$. Lemma 3.4.1 may also be used to rewrite the weights in a form that more closely resembles that of EL:

$$\omega_i(\theta) = \frac{1}{n} \left( \frac{1}{1 + \delta g_i(\theta)S_n(\theta)G_n(\theta) - \delta(1 - \delta)G_n(\theta)'S_n(\theta)W_n^{-1}S_n(\theta)G_n(\theta)} \right).$$  \hspace{1cm} (3.33)
Theorem 3.4.2 (Equivalency)

Suppose the conditions of Lemma 3.4.1 are met. Let \( \delta \) approach zero. Then PMM’s first-order conditions limit to those of GMM. Alternatively, let \( \delta \) approach one. Then PMM’s first-order conditions limit to those of EL.

Theorem 3.4.2 shows that both GMM and EL are special cases of PMM. Through \( \delta \), PMM provides a continuum of estimators with GMM and EL at the extremes. The result is obtained by applying equations (3.30), (3.31), and (3.33). When \( \delta \) approaches one, PMM’s weights limit to those of EL and \( S_n(\theta) \rightarrow V_n^{-1}(\theta) \). When \( \delta \) approaches zero, PMM’s weights limit to the fixed weights of GMM and \( S_n(\theta) \rightarrow W_n \).

Lemma 3.4.3 Suppose that all the conditions required by Lemma 3.4.1 are met. In addition, suppose that \( W_n = \Sigma^{-1} + \mathcal{O}\left(n^{-\frac{1}{2}}\right) \) is symmetric positive semi-definite. Then,

(i) \( \omega_i(\theta_0) - n^{-1} = \mathcal{O}\left(n^{-\frac{1}{2}}\right) \) \( \forall i \),

(ii) \( G_n(\theta_0) = (1 - \delta)G_n(\theta_0) + \mathcal{O}\left(n^{-1}\right) \),

(iii) \( V_n(\theta_0) = \Sigma + \mathcal{O}\left(n^{-\frac{1}{2}}\right) \),

(iv) \( M_n(\theta_0) = M_n(\theta_0) + \mathcal{O}\left(n^{-\frac{1}{2}}\right) \).

Lemma 3.4.3 describes the limiting behavior of the probability weights, the efficiently estimated moment vector, second moment matrix, and Jacobian term. The weights converge to fixed weights, and the weighted-average Jacobian and second-moment matrices converge to the limiting value of their respective equal-weight counterparts. The weighted average moment vector, on the other hand, converges to the sample average scaled by \( (1 - \delta) \). As \( \delta \) approaches one and PMM’s first-order conditions approach those of EL, \( G_n(\theta_0) \rightarrow 0 \), which is the constraint imposed by EL estimation.

Theorem 3.4.4 (Consistency)

Suppose that \( E[g_i(\theta_0)g_i(\theta_0)'] \) is positive definite, \( m_i(\theta) \) is continuous in the neighborhood of the true value \( \theta_0 \), \( \|m_i(\theta)\| \) and \( \|g_i(\theta)\|^3 \) are bounded by some integrable function \( H(x) \) in
this neighborhood, and \( E[m_i(\theta_0)] \) has rank \( k \). Then, as \( n \to \infty \), with probability 1, \( Q(\theta)_{\text{PMM}} \) attains its minimum value at some point \( \hat{\theta} \) in the interior of the ball \( \|\theta - \theta_0\| \leq n^{-\varphi} \) where \( 0 < \varphi < \frac{1}{2} \) and \( \hat{\theta} \) satisfies \( \mathcal{M}_n(\hat{\theta})'S_n(\hat{\theta})G_n(\hat{\theta}) = 0 \).

**Theorem 3.4.5 (Asymptotic Normality)**

Suppose that \( \hat{\theta} \) satisfies equation (3.29) where \( W_n \xrightarrow{p} W, W_n \) is full rank and symmetric positive semi-definite, and \( \hat{\theta} \xrightarrow{p} \theta_0 \). Let \( S \) denote the limiting value of \( S_n(\theta_0) \) as \( n \to \infty \). Further suppose that

1. \( \theta_0 \in \text{interior}(\Theta) \);
2. \( G_n(\theta) \) and \( G_n(\theta) \) are continuously differentiable in a neighborhood \( \mathcal{N} \) of \( \theta_0 \);
3. \( \sqrt{n}G_n(\theta_0) \xrightarrow{d} \mathcal{N}(0, \Sigma) \);
4. there is \( M(\theta) \) that is continuous at \( \theta_0 \) and \( \sup_{\theta \in \mathcal{N}} \|M_n(\theta) - M(\theta)\| \xrightarrow{p} 0 \);
5. there is \( \mathcal{M}(\theta) \) that is continuous at \( \theta_0 \) and \( \sup_{\theta \in \mathcal{N}} \|\mathcal{M}_n(\theta) - \mathcal{M}(\theta)\| \xrightarrow{p} 0 \);
6. for \( M = M(\theta_0), M'SM \) is nonsingular.

Then \( \sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{d} \mathcal{N}(0, (M'SM)^{-1}M'S\Sigma SM(M'SM)^{-1}) \).

Theorem 3.4.5 provides the first order asymptotic distribution of the parameter estimate for any supplied symmetric positive definite weighting matrix \( W_n \). Note the asymptotic variance is similar in form to that of non-efficient GMM, except that \( S \), a function of the weighting matrix \( W \), takes the place of \( W \) in GMM.

**Theorem 3.4.6 (Efficiency)**

Suppose that all the conditions of Theorem 3.4.5 are satisfied. Further suppose that \( W_n \xrightarrow{p} \Sigma^{-1} \). Then \( \hat{\theta} \) is efficient in that it achieves the lowest first order asymptomatic variance and \( \sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{d} \mathcal{N}(0, (M'SM)^{-1}M'S\Sigma SM(M'SM)^{-1}) \).

Theorem 3.4.6 shows that PMM has the same efficiency condition as GMM – when \( W_n \) is a consistent estimate of \( \Sigma^{-1} \) an efficient estimate of \( \hat{\theta} \) is obtained. Because a consistent estimate of \( \Sigma^{-1} \) is needed to achieve efficiency, PMM is a two-step estimator like GMM and inherits the undesired properties associated with two-step estimation. The second-step
estimate of $\theta$ depends on the first-step weighting matrix. Also, two-step estimators are not invariant to linear transformations of the vector of moment conditions.

**Theorem 3.4.7 (Test of Overidentifying Restrictions)**

Suppose that $\hat{\theta}$ satisfies equation (3.29) and $\hat{\theta} \xrightarrow{p} \theta_0$. Further suppose that $W_n \xrightarrow{p} \Sigma^{-1}$. Then

$$Q(\hat{\theta}) \xrightarrow{\text{PMM}} \chi^2_{m-k}. \quad (3.38)$$

Theorem 3.4.7 shows that the objective function value serves as a test statistic of overidentifying restrictions when the parameters are estimated efficiently. The objective function has two components. The first component is the quadratic penalty of GMM that, when appropriately scaled, is similar in structure to GMM’s test statistic of overidentifying restrictions. The second component is the KLIC penalty of EL that, when appropriately scaled, is similar in structure to EL’s likelihood ratio test statistic of overidentifying restrictions. It is possible, and asymptotically equivalent, to use the properly scaled components of PMM’s objective function as test statistics as well if so desired: $n(1-\delta)^{-2}G_n(\hat{\theta})'W_n^{-1}G_n(\hat{\theta}) \xrightarrow{d} \chi^2_{m-k}$ and $-2\delta^{-2}\sum_i \ln(n\hat{\omega}_i) \xrightarrow{d} \chi^2_{m-k}$.

**Theorem 3.4.8 (Higher-order $O(n^{-1})$ Bias)**

Suppose that $\hat{\theta}$ satisfies equation (3.29) where $W_n \xrightarrow{p} \Sigma^{-1}$, $\hat{\theta} \xrightarrow{p} \theta_0$, and $\tilde{W}$ is the preliminary first-round weighting matrix. Then

$$\text{Bias}(\hat{\theta}) = B_I + (1-\delta)B_M + (1-\delta^2)B_{\Sigma} + (1-\delta)B_{W}. \quad (3.39)$$

Theorem 3.4.8 provides the higher-order $O(n^{-1})$ bias for the family of PMM estimators. The PMM bias has a form that is similar to that of GMM and EL, is comprised of the four bias components discussed earlier, and is a continuous function of the $\delta$ parameter. As $\delta$ approaches zero, the coefficients on $B_M$, $B_{\Sigma}$, and $B_{W}$ approach one and the PMM bias approaches that of GMM. Similarly, as $\delta$ approaches one, the coefficients to $B_M$, $B_{\Sigma}$, and
$B_W$ approach zero and the PMM bias approaches that of EL.

Although the PMM higher-order bias is continuous and approaches the higher-order bias of GMM and EL in the extremes, the PMM estimate itself may be discontinuous in $\delta$ for a given sample. Consider an example in which the objective function for GMM, EL, and PMM is bimodal with two local minima. GMM and EL’s objective functions may select different local minima as the global minimum. In this case, the parameter $\delta$ will not provide a continuous transition from one local minima to the other. Instead, a more likely outcome is that a particular value of $\delta$ exists such that the two local minima have equal PMM objective function values. For this particular choice of $\delta$, the parameter vector is not identified – when $\delta$ is below this value, the optimal parameter vector is at the local minima selected by GMM to be the global minimum and when $\delta$ is above this value, the optimal parameter vector is at the local minima selected by EL to be the global minimum.

### 3.5 EL, Negative Weights, and Shrinkage

EL is one estimator within a subset of the CR estimators with an objective function that is undefined for non-positive weights. Yet, its first-order conditions may imply that, conditional on $\theta$, negative weights are optimal. For EL to be defined for a given parameter vector, the following condition must be satisfied:

$$1 + g_i(\theta)\gamma_n^{-1}(\theta)G_n(\theta) \geq 0. \quad (3.40)$$

However, the above restriction may be violated at the population parameter vector. The probability that the population parameter vector is not within the support of the parameter space decreases with sample size because $G_n(\theta_0) = O\left(n^{-\frac{1}{2}}\right)$ and increases with model misspecification because $E[G_n(\theta_0)] \neq 0$ in a misspecified model. Even when the model is correctly specified, if $\sup_{x \in \mathcal{X}} g(x, \theta_0)$ is not bounded then asymptotically there is always a positive probability that restriction (3.40) is not satisfied at the population parameter value.
Violation of condition (3.40) leads to two possible outcomes. The first scenario is that the condition is violated for a subset of the parameter space, restricting the parameter space to the region in which the condition is satisfied. The second scenario is that the restriction is violated for the entire parameter space, which occurs when the convex hull of the moments $g_1(\theta), \ldots, g_n(\theta)$ does not contain the origin for any choice of $\theta$. When this is the case, EL is undefined for the given model and sample. One possible way to implement EL estimation in the latter case is to shrink the optimal weights towards fixed weights as suggested by Bonnal and Renault (2004). They recommend that the optimal EL weights be modified in the following manner to guarantee that $\omega^*_i(\theta)_{\text{EL}} \geq 0$:

$$
\omega^*_i(\theta)_{\text{EL}} = \frac{1}{1 + \varepsilon(\theta)} \omega_i(\theta)_{\text{EL}} + \frac{\varepsilon(\theta)}{1 + \varepsilon(\theta)} \frac{1}{n}
$$  \hspace{1cm} (3.41)

$$
\varepsilon(\theta) = -n \min \left( \min_i \omega_i(\theta)_{\text{EL}}, 0 \right).
$$  \hspace{1cm} (3.42)

The optimization then proceeds with the $n - 1$ positive weights.

PMM, through $\delta$, provides the econometrician with a similar solution when EL is undefined. As described by Theorem 3.4.2, when $\delta$ approaches one, PMM’s behavior limits to that of EL and as $\delta$ approaches zero, PMM’s behavior approaches that of GMM. Like EL, PMM also has an objective function that disallows non-positive weights. For PMM to be defined for a given parameter vector, the following condition must be satisfied:

$$
1 + \delta g_i(\theta)^t S_n(\theta) G_n(\theta) - \delta(1 - \delta) G_n(\theta)^t S_n(\theta) W_n^{-1} S_n(\theta) G_n(\theta) \geq 0.
$$  \hspace{1cm} (3.43)

The coefficient $\delta$ may be selected to ensure that positive weights are attainable for some region of the parameter space. There are two differences between shrinking through PMM vs. the Bonnal and Renault (2004) approach. The first difference is that, for PMM, an equal amount of shrinkage occurs throughout the parameter space, whereas the Bonnal and Renault (2004) shrinkage is conditional on $\theta$. The second difference is that Bonnal and Renault (2004) shrink the optimal weights towards fixed weights, whereas PMM selects optimal weights as a function of the shrinkage imposed by $\delta$. 
It may be desirable to implement PMM estimation to increase the support of the parameter space even when the parameter space support under EL is non-empty. Suppose that for two parameter vectors, $\theta_1$ and $\theta_2$,

$$nG_n(\theta_1) \nu_n^{-1}(\theta_1)G_n(\theta_1) < nG_n(\theta_2) \nu_n^{-1}(\theta_2)G_n(\theta_2).$$

Then for the given data set, the model is better specified under the first parameter vector. It is possible, however, that for one or more observations,

$$g_i(\theta_1) \nu_n^{-1}(\theta_1)G_n(\theta_1) < -1$$

and

$$g_i(\theta_2) \nu_n^{-1}(\theta_2)G_n(\theta_2) \geq -1.$$  

Then, EL is undefined for the better specified model and selects $\theta_2$ in order to avoid the negative weight(s). For reasons described above, this hypothetical scenario is more likely to occur when the sample size is small, the number of moments is large, the moments are unbounded, and the model is misspecified. By shrinking the weights through $\delta$, PMM is able to offer a less restricted parameter space than EL. We may select $\delta$ such that the optimal weights are positive for both $\theta_1$ and $\theta_2$ and the estimator will then pick $\theta_1$ over $\theta_2$ as desired.

The natural question is, why not let $\delta$ approach zero and avoid the scenario outlined in the previous paragraph entirely? The answer is that doing so would prohibit us from attaining the desirable higher-order properties offered by EL. PMM allows the econometrician to exchange the higher-order properties provided by EL for the less restricted parameter space associated with GMM and vice-versa.
3.6 Computation

At first glance, EL may appear to have a computational advantage over PMM. Unlike the GEL estimators, where the \( n \)-dimensional weighting vector is a function of an \( m \)-dimensional vector, the PMM weighting vector is the solution to the \( n \) equations given by (3.28). Here, I show how to redefine the problem to achieve a mapping from \( \mathbb{R}^m \) to \( \mathbb{R}^n \). In so doing, the computational complexity of PMM is reduced to that of the GEL estimators.

Rewrite equation (3.27) replacing \( G_n(\theta) \) with the \( m \)-dimensional vector \( \lambda \):

\[
\omega_i(\lambda, \theta) = \frac{1}{N} \left( \frac{1}{1 + \frac{\delta}{1-\delta}(g_i(\theta) - \lambda)'W_n\lambda} \right). \tag{3.44}
\]

The objective is to find the \( \lambda \) such that \( \lambda \) is indeed \( G_n(\theta) \). Define

\[
f(\lambda, \theta) = \lambda - \sum_i \omega_i(\lambda, \theta)g_i(\theta). \tag{3.45}
\]

Then \( \lambda \) may be obtained via the Newton-Raphson iterative algorithm:

\[
\lambda_{j+1} = \lambda_j - \left[ \mathbb{I}_m + n \frac{\delta}{1-\delta} \sum_i \omega_i^2(\lambda_j, \theta)g_i(\theta) (g_i(\theta) - 2\lambda_j)'W_n \right]^{-1} \left[ \lambda_j - \sum_i \omega_i(\lambda_j, \theta)g_i(\theta) \right]. \tag{3.46}
\]

Lemma 3.4.3 suggests an appropriate starting value for the iteration: \( \lambda_0 = (1 - \delta)G_n(\theta) \).

For my simulations, I find that setting \( \lambda_0 = 0.25(1 - \delta)G_n(\theta) \) leads to fewer instances of divergence. The iterative procedure described above constitutes an inner loop which must be solved for any given \( \theta \) yielding \( \lambda(\theta) \). The outer loop selects \( \theta \) to minimize the objective function defined by equation (3.22):

\[
\hat{\theta} = \arg\min_{\theta \in \Theta} \frac{1}{1-\delta} n\lambda(\theta)'W_n\lambda(\theta) + \frac{2}{\delta} \sum_i \ln \left( 1 + \frac{\delta}{1-\delta}(g_i(\theta) - \lambda(\theta))'W_n\lambda(\theta) \right). \tag{3.47}
\]
3.7 Monte Carlo Simulations

In this section, I compare the specific PMM estimator obtained by setting $\delta = 0.5$ to EL through a Monte Carlo study. The goal of the simulations is to determine what type of improvement, if any, is achieved by expanding the parameter space through shrinkage. If the scenario detailed in an earlier section is an issue, we would expect the PMM estimator to have test statistics that are less mis-sized when the number of observations is low or the number of moments is high. In the experiments conducted below, I compare PMM to EL across several dimensions. I calculate the realized estimator mean, volatility, and Root Mean Squared Error (RMSE). In addition, I provide the size of the rejection region for the test of overidentifying restriction with $\alpha$ chosen to be 0.10, 0.05, and 0.01 percent. Finally, I calculate the probability of rejecting the population parameter value for $\alpha$ equal to 0.10, 0.05, and 0.01 percent.

I simulate the model proposed by Hall and Horowitz (1996) as modified and investigated by Schennach (2006), who expands the number of moment conditions from 2 to $K$. The orthogonality conditions are

$$g_i(\theta) = r(x_i, \theta) [1 \ x_{i2} \ (x_{i3} - 1) \ \ldots \ (x_{iK} - 1)]', \quad (3.48)$$

where $r(x_i, \theta) = \exp(-0.72 - (x_{i1} + x_{i2})\theta + 3x_{i2}) - 1$. When $(x_{i1}, x_{i2}) \sim N(0, 0.16I_2)$ and $x_{ik} \sim \chi^2_1$, for $k = 3, \ldots, K$ the moment conditions are satisfied at $\theta_0 = 3$. Because the third moments of all elements of $g_i(\theta)$ are non-zero and $g_i(\theta)$ is a nonlinear function of $\theta$, the $O(n^{-1})$ term does not trivially vanish.

I perform simulations with sample sizes 25, 50, and 100, because negative weights are more likely to be occur under small samples. For the 25 sample study, I consider a 2 moment and 5 moment model. For the 50 sample study, I simulate with 2, 5, and 10 moments. The simulations with sample size 100 are conducted with 2, 5, 10, 15, and 20 moments. Each batch of simulations includes 2000 trials. The results of the simulations are presented in both tabular and graphical format. Table 1 reports the results for the 25 and 50 sample estimations and Table 2 reports the results for the 100 sample estimations. For each Monte
Carlo experiment, I plot the densities of the realized estimates obtained under EL and PMM estimation. Figures 1 and 2 provide the plots for the 25 observation simulations. Figures 3 through 5 plot the densities for the 50 observation experiments. The densities for estimates obtained under sample size of 100 are plotted in Figures 6 through 10.

When the sample size is small and the number of moments are low, EL and PMM have similar properties. For the 25 observation and 2 moment condition trials, the realized volatility of the EL and PMM estimators are 0.78 and 0.79 respectively. Hypothesis tests for the parameter are equally mis-sized. For $\alpha = 0.05$, the EL estimate is rejected 40 percent of the time and the PMM estimate is rejected 39 percent of the time. For tests of model-specification, EL is slightly less mis-sized. For $\alpha = 0.05$, the model is rejected 12 percent of the time for EL and 15 percent of the time for PMM. For the 50 sample simulations, the results are similar under the models with 2 and 5 moment restrictions. With 2 moment conditions, EL’s estimates are less volatile: 0.42 for EL and 0.49 for PMM. With 5 moment conditions, the estimator volatilities are almost equal: 0.46 for EL and 0.47 for PMM. Both estimators have roughly equivalent sized hypothesis tests. For 50 observations and 2 moments, EL rejects the population parameter at the $\alpha = 0.05$ level 28 percent of the time whereas PMM rejects 31 percent of the time.

Significant differences between the two estimators occur under the models with large number of moments. With 25 observations and 5 moments, EL’s estimator volatility is approximately 0.94 while PMM’s estimator volatility is 0.85. At the 5 percent level, PMM rejects the population parameter value 36 percent of the time, which EL rejects 44 percent of the time. PMM and EL reject the model at the 5 percent level, respectively, 61 and 48 percent of the time. The results are similar for the 50 observation and 10 moment model. EL’s estimator volatility is 0.724, which is approximately 18 percent higher than that of PMM. PMM and EL reject the population parameter value at the 5 percent level approximately 26 and 43 percent of the time respectively. The model is rejected at the 5 percent level 63 percent of the time for PMM and 73 percent of the time for EL.

For the simulations with 100 observations, PMM and EL have similar properties when there are 10 moment conditions. PMM’s estimator volatility is 0.33, whereas EL’s is 0.32. PMM
and EL reject the model at the 5 percent level 59 and 61 percent of the time respectively and the two estimators reject the population parameter value 21 and 16 percent of the time respectively. EL’s estimates are less volatile than PMM when the number of moments are low, but its volatility increases more than PMM as the number of moment conditions grows. EL’s estimator volatility increases from 0.30 under 2 moments to 0.45 under 20 moments. PMM’s estimator, on the other hand, increases from 0.32 under 2 moments to 0.37 with 20 moment conditions.

The differences in estimate volatilities reported in the tables and described above are visually apparent in the figures. In Figures 1, 3, 4, 6, and 7, which represent the 25x2, 50x2, 50x5, 100x2, and 100x5 simulations, the densities of EL and PMM’s estimates are almost identical. In Figures 2, 5, 9, and 10, which represent the 25x5, 50x5, 100x15, and 100x20 trials, PMM’s density functions have higher peaks and tighter distributions than those of EL. Finally, Figure 8 displays the density for the 100 observation and 10 moment conditions experiment and shows EL to have the denser distribution.

The series of experiments suggest that when the number of observations is large or the number of moments conditions is low, EL and PMM have similar properties; EL’s estimates are slightly less volatile and its hypothesis tests tend to be less mis-sized. When the number of moments conditions is large, particularly when the number of observations is low, the restricted parameter space due to negative weight avoidance leads EL to have more volatile estimates and greater over-rejection of hypothesis tests. Shrinking the weights towards fixed weights through PMM appears to provide more desirable estimator properties.

3.8 Conclusion

This paper presents a new family of estimators by merging the objective functions of GMM and EL. Each of the two estimators have opposite sources of rigidity and flexibility. GMM allows the sample average moment vector to be non-zero and requires the sample weights to be fixed. EL allows the sample weights to vary and requires the weighted moment vector to equal zero. By merging the two objective functions, the PMM family of estimators allows
both weights and sample moments to deviate from $n^{-1}$ and 0 respectively and measures the respective deviations with EL’s KLIC penalty function and GMM’s quadratic penalty function. Through a free parameter, a continuum of estimators is obtained with GMM and EL at the extremes.

EL’s objective function is undefined for non-positive weights. This property restricts the parameter space to the set of parameters that result in positive optimal weights. The population parameter vector may not fall within the restricted parameter space, an event that is more likely when the sample size is small, the number of moments is large, and the model is misspecified. PMM provides a mechanism in which negative weights may be shrunk towards $n^{-1}$, leading to a less restricted parameter space.

By performing a series of Monte Carlo experiments for the Hall and Horowitz (1996) model as modified by Schennach (2006) I compare the EL and PMM estimators under conditions where the restricted parameter space is likely to be problematic. The main findings are that when the sample size is large or the number of moments are low, the EL and PMM estimates have similar properties. The estimate volatilities and the size of hypothesis tests are approximately the same. When the sample size is low and the number of moments are large, however, shrinking the weights through PMM appears to provide estimates with more desirable properties. PMM’s estimate volatilities are lower and its hypothesis tests are less mis-sized than EL under these conditions.

PMM is a continuum of estimators with GMM and EL at the extremes, but the simulations only consider the PMM estimator that is ‘halfway’ between GMM and EL. I suspect additional work on ‘optimal’ shrinkage will prove valuable. My conjecture is that models with smaller samples and larger moments will benefit from greater shrinkage towards GMM, while those with larger samples and less moment conditions will have more desirable properties when PMM is selected to more closely resemble EL.
Table 3.1: Monte Carlo Simulations. This table presents the results of the Monte Carlo study for the Hall and Horowitz (1996) model as modified by Schennach (2006) with sample sizes of 25 and 50 observations and 2000 trials. The population parameter value $\theta_0$ is 3. $E(\hat{\theta})$ is the sample average of the estimates, $\sigma(\hat{\theta})$ is the sample volatility of the estimates, and $rmse(\hat{\theta})$ is the Root Mean Squared Error of the estimates. $\chi^2_{0.10}$ and $t_{0.05}$ respectively represent the size of the rejection region for the test of overidentifying restrictions and test that the estimate equals the population parameter value at the $\alpha$ level.

<table>
<thead>
<tr>
<th>Model</th>
<th>$E(\hat{\theta})$</th>
<th>$\sigma(\hat{\theta})$</th>
<th>$rmse(\hat{\theta})$</th>
<th>$\chi^2_{0.10}$</th>
<th>$\chi^2_{0.05}$</th>
<th>$t_{0.10}$</th>
<th>$t_{0.05}$</th>
<th>$t_{0.01}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>EL</td>
<td>3.092 (0.017)</td>
<td>0.778 (0.020)</td>
<td>0.783 (0.020)</td>
<td>0.199 (0.009)</td>
<td>0.121 (0.007)</td>
<td>0.036 (0.004)</td>
<td>0.444 (0.011)</td>
<td>0.396 (0.011)</td>
</tr>
<tr>
<td>PMM</td>
<td>3.110 (0.018)</td>
<td>0.791 (0.021)</td>
<td>0.798 (0.021)</td>
<td>0.223 (0.009)</td>
<td>0.154 (0.008)</td>
<td>0.054 (0.005)</td>
<td>0.443 (0.011)</td>
<td>0.392 (0.011)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>25 Observations and 5 Moment Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>EL</td>
</tr>
<tr>
<td>PMM</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>50 Observations and 2 Moment Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>EL</td>
</tr>
<tr>
<td>PMM</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>50 Observations and 5 Moment Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>EL</td>
</tr>
<tr>
<td>PMM</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>50 Observations and 10 Moment Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>EL</td>
</tr>
<tr>
<td>PMM</td>
</tr>
</tbody>
</table>
Table 3.2: Monte Carlo Simulations. This table presents the results of the Monte Carlo study for the Hall and Horowitz (1996) model as modified by Schennach (2006) with sample size of 100 observations and 2000 trials. The population parameter value $\theta_0$ is 3. $E(\hat{\theta})$ is the sample average of the estimates, $\sigma(\hat{\theta})$ is the sample volatility of the estimates, and $rmse(\hat{\theta})$ is the Root Mean Squared Error of the estimates. $\chi^2_{0.10}$ and $t_{0.01}$ respectively represent the size of the rejection region for the test of overidentifying restrictions and test that the estimate equals the population parameter value at the $\alpha$ level.

<table>
<thead>
<tr>
<th>Model</th>
<th>$E(\hat{\theta})$</th>
<th>$\sigma(\hat{\theta})$</th>
<th>$rmse(\hat{\theta})$</th>
<th>$\chi^2_{0.10}$</th>
<th>$\chi^2_{0.05}$</th>
<th>$t_{0.10}$</th>
<th>$t_{0.05}$</th>
<th>$t_{0.01}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>EL</td>
<td>3.058 (0.007)</td>
<td>0.299 (0.005)</td>
<td>0.304 (0.006)</td>
<td>0.186 (0.009)</td>
<td>0.119 (0.007)</td>
<td>0.039 (0.004)</td>
<td>0.271 (0.010)</td>
<td>0.222 (0.009)</td>
</tr>
<tr>
<td>PMM</td>
<td>3.050 (0.007)</td>
<td>0.319 (0.007)</td>
<td>0.323 (0.007)</td>
<td>0.182 (0.009)</td>
<td>0.114 (0.007)</td>
<td>0.038 (0.004)</td>
<td>0.284 (0.010)</td>
<td>0.232 (0.009)</td>
</tr>
</tbody>
</table>

100 Observations and 5 Moment Conditions

| EL    | 3.148 (0.007)    | 0.322 (0.007)    | 0.354 (0.008)   | 0.404 (0.011)   | 0.311 (0.010)   | 0.169 (0.008) | 0.297 (0.010) | 0.246 (0.010) | 0.165 (0.008) |
| PMM   | 3.093 (0.007)    | 0.318 (0.007)    | 0.331 (0.007)   | 0.433 (0.011)   | 0.325 (0.010)   | 0.170 (0.008) | 0.265 (0.010) | 0.210 (0.009) | 0.139 (0.008) |

100 Observations and 10 Moment Conditions

| EL    | 3.054 (0.007)    | 0.320 (0.010)    | 0.324 (0.010)   | 0.700 (0.011)   | 0.605 (0.011)   | 0.413 (0.011) | 0.214 (0.009) | 0.162 (0.008) | 0.107 (0.007) |
| PMM   | 3.144 (0.006)    | 0.327 (0.007)    | 0.357 (0.007)   | 0.679 (0.010)   | 0.592 (0.011)   | 0.402 (0.011) | 0.267 (0.010) | 0.210 (0.009) | 0.144 (0.008) |

100 Observations and 15 Moment Conditions

| EL    | 3.376 (0.009)    | 0.412 (0.009)    | 0.558 (0.010)   | 0.846 (0.008)   | 0.783 (0.009)   | 0.642 (0.011) | 0.440 (0.011) | 0.381 (0.011) | 0.296 (0.010) |
| PMM   | 3.001 (0.008)    | 0.340 (0.008)    | 0.340 (0.008)   | 0.848 (0.008)   | 0.776 (0.008)   | 0.581 (0.011) | 0.189 (0.009) | 0.143 (0.009) | 0.084 (0.006) |

100 Observations and 20 Moment Conditions

| EL    | 3.442 (0.010)    | 0.454 (0.011)    | 0.634 (0.012)   | 0.935 (0.006)   | 0.895 (0.007)   | 0.805 (0.009) | 0.469 (0.011) | 0.403 (0.011) | 0.313 (0.010) |
| PMM   | 2.885 (0.008)    | 0.370 (0.008)    | 0.387 (0.009)   | 0.919 (0.006)   | 0.871 (0.007)   | 0.712 (0.010) | 0.177 (0.009) | 0.125 (0.007) | 0.076 (0.006) |
Figure 3.1: **Kernel Density Plot for 25 Observations and 2 Moment Conditions.** This figure plots the kernel density of the EL and PMM estimates obtained for a Monte Carlo study for the Hall and Horowitz (1996) model as modified by Schennach (2006) with sample size of 25 observations, 2 moment conditions, and 2000 trials. The population parameter value is $\theta_0 = 3$. PMM is represented by the bold line.
Figure 3.2: **Kernel Density Plot for 25 Observations and 5 Moment Conditions.** This figure plots the kernel density of the EL and PMM estimates obtained for a Monte Carlo study for the Hall and Horowitz (1996) model as modified by Schennach (2006) with sample size of 25 observations, 5 moment conditions, and 2000 trials. The population parameter value is $\theta_0 = 3$. PMM is represented by the bold line.
Figure 3.3: **Kernel Density Plot for 50 Observations and 2 Moment Conditions.** This figure plots the kernel density of the EL and PMM estimates obtained for a Monte Carlo study for the Hall and Horowitz (1996) model as modified by Schennach (2006) with sample size of 25 observations, 2 moment conditions, and 2000 trials. The population parameter value is $\theta_0 = 3$. PMM is represented by the bold line.
Figure 3.4: **Kernel Density Plot for 50 Observations and 5 Moment Conditions.** This figure plots the kernel density of the EL and PMM estimates obtained for a Monte Carlo study for the Hall and Horowitz (1996) model as modified by Schennach (2006) with sample size of 50 observations, 5 moment conditions, and 2000 trials. The population parameter value is $\theta_0 = 3$. PMM is represented by the bold line.
Figure 3.5: **Kernel Density Plot for 50 Observations and 10 Moment Conditions.** This figure plots the kernel density of the EL and PMM estimates obtained for a Monte Carlo study for the Hall and Horowitz (1996) model as modified by Schennach (2006) with sample size of 50 observations, 10 moment conditions, and 2000 trials. The population parameter value is $\theta_0 = 3$. PMM is represented by the bold line.
Figure 3.6: **Kernel Density Plot for 100 Observations and 2 Moment Conditions.** This figure plots the kernel density of the EL and PMM estimates obtained for a Monte Carlo study for the Hall and Horowitz (1996) model as modified by Schennach (2006) with sample size of 100 observations, 2 moment conditions, and 2000 trials. The population parameter value is $\theta_0 = 3$. PMM is represented by the bold line.
Figure 3.7: **Kernel Density Plot for 100 Observations and 5 Moment Conditions.** This figure plots the kernel density of the EL and PMM estimates obtained for a Monte Carlo study for the Hall and Horowitz (1996) model as modified by Schemach (2006) with sample size of 100 observations, 5 moment conditions, and 2000 trials. The population parameter value is $\theta_0 = 3$. PMM is represented by the bold line.
Figure 3.8: Kernel Density Plot for 100 Observations and 10 Moment Conditions. This figure plots the kernel density of the EL and PMM estimates obtained for a Monte Carlo study for the Hall and Horowitz (1996) model as modified by Schennach (2006) with sample size of 100 observations, 10 moment conditions, and 2000 trials. The population parameter value is $\theta_0 = 3$. PMM is represented by the bold line.
Figure 3.9: **Kernel Density Plot for 100 Observations and 15 Moment Conditions.** This figure plots the kernel density of the EL and PMM estimates obtained for a Monte Carlo study for the Hall and Horowitz (1996) model as modified by Schemach (2006) with sample size of 100 observations, 15 moment conditions, and 2000 trials. The population parameter value is $\theta_0 = 3$. PMM is represented by the bold line.
Figure 3.10: **Kernel Density Plot for 100 Observations and 20 Moment Conditions.** This figure plots the kernel density of the EL and PMM estimates obtained for a Monte Carlo study for the Hall and Horowitz (1996) model as modified by Schemnach (2006) with sample size of 100 observations, 20 moment conditions, and 2000 trials. The population parameter value is \( \theta_0 = 3 \). PMM is represented by the bold line.
Bibliography


Glosten, L. R., and P. R. Milgrom (1985): “Bid, Ask and Transaction Prices in a Spe-


Future Liquidity, Present Value: Measuring and Pricing Liquidity Risk

The Jackknife

Beginning with the full sample of \( T \) observations and their respective lags, \( T \) new resampled data sets with \( T - 1 \) observations each are created by removing a different single observation from the original dataset. The first resampled data set will have the first observation removed, the second will have the second observation removed, and so on. The VAR is then re-estimated for each resampled data set. For example, the following estimation would be performed for the second resampling:

\[
\begin{bmatrix}
  z_1 \\
  z_2 \\
  z_3 \\
  : \\
  z_T
\end{bmatrix}
= \begin{bmatrix}
  z_0 \\
  z_1 \\
  z_2 \\
  : \\
  z_{T-1}
\end{bmatrix}
+ \begin{bmatrix}
  w_1 \\
  w_2 \\
  w_3 \\
  : \\
  w_T
\end{bmatrix} \Gamma_2 + \begin{bmatrix}
  w_1 \\
  w_2 \\
  w_3 \\
  : \\
  w_T
\end{bmatrix} \Sigma_2 + \begin{bmatrix}
  w_1 \\
  w_2 \\
  w_3 \\
  : \\
  w_T
\end{bmatrix}.
\] (49)

For any statistic of interest, this procedure yields a total of \( T + 1 \) estimates: the original estimate \( f(\Gamma, \Sigma) \) and the \( T \) resampled estimates, \( f(\Gamma_i, \Sigma_i) \), where \( \Gamma_i \) and \( \Sigma_i \) are the estimates of the parameter matrix and error-covariance matrix for each resampling. The standard
error of the statistic, \( \sigma(f(\Gamma, \Sigma)) \), is calculated according to the formula

\[
\sigma(f(\Gamma, \Sigma)) = \left[ \frac{T - 1}{T} \sum_{i=1}^{T} (f(\Gamma_i, \Sigma_i) - \bar{f}(\Gamma, \Sigma))^2 \right]^{\frac{1}{2}},
\]

where \( \bar{f}(\Gamma, \Sigma) \) is the mean of the \( T \) resampled statistics

\[
\bar{f}(\Gamma, \Sigma) = \frac{1}{T} \sum_{i=1}^{T} f(\Gamma_i, \Sigma_i).
\]

In addition to providing an estimate of robust standard errors, the jackknife may be used to adjust for small-sample bias. The small-sample corrected estimate of a statistic, \( \hat{f}(\Gamma, \Sigma) \) is calculated according to the formula:

\[
\hat{f}(\Gamma, \Sigma) = T f(\Gamma, \Sigma) - (T - 1) \bar{f}(\Gamma, \Sigma).
\]

The Acharya and Pedersen (2005) Betas and Mean Reversion

Let \( \tau \) denote the holding period of an asset in months and \( r_{t,t+\tau}^i \) denote the log gross return for asset \( i \) over the \( \tau \) month period between time \( t \) and \( t + \tau \):

\[
r_{t,t+\tau} = r_t + r_{t+1} + \cdots + r_{t+\tau-1} + r_{t+\tau}
\]

The four betas in the Acharya and Pedersen (2005) adjusted CAPM are:

\[
\beta_1 = \frac{ \text{cov}_t(r_{t,t+\tau}^i, r_{t,t+\tau}^m) }{ \text{var}_t(r_{t,t+\tau}^m - K_{t+\tau}^m) }
\]

\[
\beta_2 = \frac{ \text{cov}_t(K_{t+\tau}^i, K_{t+\tau}^m) }{ \text{var}_t(r_{t,t+\tau}^m - K_{t+\tau}^m) }
\]

\[
\beta_3 = \frac{ \text{cov}_t(r_{t,t+\tau}^i, K_{t+\tau}^m) }{ \text{var}_t(r_{t,t+\tau}^m - K_{t+\tau}^m) }
\]

\[
\beta_4 = \frac{ \text{cov}_t(K_{t+\tau}^i, r_{t,t+\tau}^m) }{ \text{var}_t(r_{t,t+\tau}^m - K_{t+\tau}^m) }
\]

Acharya and Pedersen (2005) estimate the four betas by assuming that returns are independent over time and that liquidity follows a random walk. Under these two assumptions,
The five covariances needed to calculate the four liquidity betas are calculated to be:

\[
\begin{align*}
\beta_1 &= \frac{\text{cov}_t(r_{t+1}^i, r_{t+1}^m)}{\text{var}_t(r_{t+1}^m - K_{t+1}^m)} \\
\beta_2 &= \frac{\text{cov}_t(K_{t+1}^i, K_{t+1}^m)}{\text{var}_t(r_{t+1}^m - K_{t+1}^m)} \\
\beta_3 &= \frac{\text{cov}_t(r_{t+1}^i, K_{t+1}^m)}{\text{var}_t(r_{t+1}^m - K_{t+1}^m)} \\
\beta_4 &= \frac{\text{cov}_t(K_{t+1}^i, r_{t+1}^m)}{\text{var}_t(r_{t+1}^m - K_{t+1}^m)}.
\end{align*}
\]

Here, I show that small deviations from the two assumptions may have a substantial impact on the estimated betas. Suppose returns and proportional costs are first order autoregressive:

\[
\begin{align*}
    r_t^i &= \alpha_t + \rho_t r_{t-1}^i + \epsilon_t^i \\
    K_t^i &= \delta_t + \varphi_t K_{t-1}^i + \epsilon_t^i.
\end{align*}
\]

Let rand\(_t(\tau\tau)\) denote the stochastic component of \(x_{t\tau}\) conditional on information at time \(t\). Then it can be shown that

\[
\begin{align*}
\text{rand}_t(K_t^i) &= \sum_{j=1}^{\tau} \varrho^{\tau-j} \epsilon_{t+j} \\
\text{rand}_t(r_t^i) &= \frac{1}{1 - \rho_t} \sum_{j=1}^{\tau} \epsilon_{t+j} - \frac{\rho_t}{1 - \rho_t} \sum_{j=1}^{\tau} \varrho^{\tau-j} \epsilon_{t+j}
\end{align*}
\]

The five covariances needed to calculate the four liquidity betas are calculated to be:

\[
\begin{align*}
\text{cov}_t(K_t^i, K_{t+1}^m) &= \frac{1 - \varrho_t^m \varrho_t^r}{1 - \varrho_t^m} \text{cov}_t(K_{t+1}^i, K_{t+1}^m) \\
\text{cov}_t(r_t^i, r_{t+1}^m) &= \frac{1}{(1 - \rho_t)(1 - \rho_m)} \left( \tau + \rho_t \rho_m \frac{1 - \rho_t^m \rho_t^\tau}{1 - \rho_t \rho_m} - \rho_t \frac{1 - \rho_t^\tau}{1 - \rho_t} \right) \text{cov}_t(r_{t+1}^i, r_{t+1}^m) \\
\text{cov}_t(K_t^i, r_{t+1}^m) &= \left( \frac{1}{1 - \rho_t \rho_m} - \frac{\rho_t \varrho_t^\tau}{1 - \rho_t \rho_m} \right) \text{cov}_t(K_{t+1}^i, r_{t+1}^m) \\
\text{cov}_t(r_t^i, K_{t+1}^m) &= \left( \frac{1}{1 - \rho_t \rho_m} - \frac{\rho_t \varrho_t^\tau}{1 - \rho_t \rho_m} \right) \text{cov}_t(r_{t+1}^i, K_{t+1}^m) \\
\text{var}_t(r_{t+1}^m - K_{t+1}^m) &= \frac{1}{(1 - \rho_m)^2} \left( \tau + \rho_m \frac{1 - \rho_m^2}{1 - \rho_m^2} - 2 \rho_m \frac{1 - \rho_m^\tau}{1 - \rho_m} \right) \text{var}_t(r_{t+1}^m) + \frac{1 - \varrho_m^2}{1 - \varrho_m} \text{cov}_t(r_{t+1}^m, K_{t+1}^m) \\
&\quad - 2 \left( \frac{1}{1 - \rho_m} - \frac{\rho_m}{1 - \rho_m \varrho_m \rho_m} \right) \text{cov}_t(r_{t+1}^m, K_{t+1}^m).
\end{align*}
\]
As an example, I consider the equal-weight market portfolio over the period 1964 – 2002. For the portfolio I estimate the AR(1) coefficients to be $\rho_m = 0.167$ and $\varrho = 0.886$ and the covariance of the residuals to be $\text{var}_t(r_{t+1}^m) = 28.6\%$, $\text{var}_t(K_{t+1}^m) = 0.033\%$, and $\text{cov}_t(r_{t+1}^m, K_{t+1}^m) = -0.397\%$. Letting $\beta_1^*, \beta_2^*, \beta_3^*$, and $\beta_4^*$ be the respective betas calculated using the first-order autoregressive assumptions and letting the AR coefficients for the individual portfolio equal that of the market portfolio, I calculate that $\beta_1 = 0.98\beta_1^*$, $\beta_2 = 6.31\beta_2^*$, $\beta_3 = 3.08\beta_3^*$, and $\beta_4 = 3.08\beta_4^*$. For this example, we see that although the calculated return beta $\beta_1$ appears to be relatively unaffected but the three liquidity betas are substantially overestimated when not taking into account their autoregressive properties. For instance, the beta that represents the systematic risk due to the comovement of a portfolio’s proportional cost with that of the market is overestimated by a factor of 6.3.
Surprise! It’s Illiquid: The January 2006 Tokyo Stock Exchange Shutdown as a Natural Experiment

Standard Error of Randomly-Formed Portfolio

Here, I derive the standard error of a portfolio that is formed by purchasing and selling equal dollar amounts of, respectively, $N_L$ and $N_S$ randomly selected assets. Further, the portfolio is created by purchasing and selling, respectively, $N_{L,1}, \ldots, N_{L,I}$ and $N_{S,1}, \ldots, N_{S,I}$ assets from $I$ industries, with $N_{L,1} + \cdots + N_{L,I} = N_L$, $N_{S,1} + \cdots + N_{S,I} = N_S$. Finally, I assume that within each industry group, all asset returns are drawn from the same distribution with variance $\sigma_i^2$ for industry $i$. Let $r_{L,n}$ and $r_{S,n}$ denote the return for asset $n$ within the respective long and short portfolios. Also, let $r_{L,n}^i$ and $r_{S,n}^i$ denote the return for asset $n$ in industry $i$ within the respective long and short portfolios. Then,

$$r_{rand} = \frac{1}{N_L} \sum_{n=1}^{N_L} r_{L,n} - \frac{1}{N_S} \sum_{n=1}^{N_S} r_{S,n}$$

$$= \frac{1}{N_{L,1}} \sum_{n=1}^{N_{L,1}} r_{L,n}^1 + \cdots + \frac{1}{N_{L,I}} \sum_{n=1}^{N_{L,I}} r_{L,n}^I - \frac{1}{N_{S,1}} \sum_{n=1}^{N_{S,1}} r_{S,n}^1 - \cdots - \frac{1}{N_{S,I}} \sum_{n=1}^{N_{S,I}} r_{S,n}^I$$

(56)
and
\[
\text{var} \left( r^{\text{rand}} \right) = \frac{N_{L,1}}{N_L^2} \sigma_1^2 + \cdots + \frac{N_{L,i}}{N_L^2} \sigma_i^2 + \frac{N_{S,1}}{N_S^2} \sigma_1^2 + \cdots + \frac{N_{S,i}}{N_S^2} \sigma_i^2
\]
\[
= \sum_{i=1}^{I} \frac{N_{L,i} N_S^2 + N_{S,i} N_L^2}{N_L^2 N_S^2} \sigma_i^2. \tag{57}
\]

When the variance of asset returns are the same across all industries, then the variance of the long–short portfolio simplifies to
\[
\text{var} \left( r^{\text{rand}} \right) = \frac{N_L + N_S}{N_L N_S} \sigma^2 \tag{58}
\]

Equation (57) may be rewritten:
\[
\text{var} \left( r^{\text{rand}} \right) = \frac{1}{N_L} \sum_{i=1}^{I} \frac{N_{L,i}}{N_L} \sigma_i^2 + \frac{1}{N_S} \sum_{i=1}^{I} \frac{N_{S,i}}{N_S} \sigma_i^2 \tag{59}
\]

If \( \frac{N_{L,i}}{N_L} \approx \frac{N_i}{N} \approx \frac{N_{S,i}}{N_S} \) and \( \sigma^2 = \sum_{i=1}^{I} \frac{N_i}{N} \sigma_i^2 \) then
\[
\text{var} \left( r^{\text{rand}} \right) \approx \frac{N_L + N_S}{N_L N_S} \sigma^2. \tag{60}
\]

The Hasbrouck (2006) Bid-Ask Spread Estimator

Hasbrouck (2006) suggests a Bayesian variant of the Roll (1984) bid-ask spread estimator. The price dynamics are:
\[
m_t = m_{t-1} + u_t \tag{61}
\]
\[
p_t = m_t + c q_t,
\]

where \( m_t \) is the log quote midpoint, \( p_t \) is the log trade price, \( c \) is half the bid-ask spread, and \( q_t \) are trade indicators with values +1 for a market buy at the ask price and −1 for a market sell at the bid price. The shock \( u_t \) is new information available and is assumed to
be uncorrelated with \( q_t \). Taking first differences of the log price provides

\[ \Delta p_t = c \Delta q_t + u_t. \tag{62} \]

If the trade indicators \( q_t \) were known, then \( c \) could be estimated using OLS. With unknown trade indicators, the Roll model considers the first-order autocovariance of log price changes:

\[ \hat{c}_{\text{Roll}} = \sqrt{-\text{cov}(\Delta p_t, \Delta p_{t-1})}, \tag{63} \]

subject to non-negativity of the autocovariance term.

The Hasbrouck (2006) specification is as follows. First, the distribution of the disturbances is assumed to be i.i.d. normal with volatility \( \sigma_u \). The data sample is simply the time-series of log prices: \( p = \{p_0, p_1, \ldots, p_T\} \). The unknowns are the model parameters \( c \) and \( \sigma_u \) and the trade direction indicators \( q = \{q_0, q_1, \ldots, q_T\} \).

When the prior distribution for \( c \) is \( \hat{c} \sim N(\hat{\mu}_c, \hat{\sigma}_c^2) \), the posterior for \( \hat{c} \) is \( \hat{c} \sim N(\hat{\mu}_c, \hat{\sigma}_c^2) \), where

\[
\hat{\mu}_c = \frac{T \cdot \text{cov}(\Delta p_t, \Delta q_t) + \frac{\sigma_u^2}{\sigma^2} \hat{\mu}_c}{T \cdot \text{var}(\Delta q_t) + \frac{\sigma_u^2}{\sigma^2}} \quad \text{and} \quad \hat{\sigma}_c^2 = \frac{\sigma_u^2}{T \cdot \text{var}(\Delta q_t) + \frac{\sigma_u^2}{\sigma^2}}. \tag{64} \]

If the prior is drawn from a truncated normal distribution, then the posterior distribution is truncated to the same interval and the parameters are unaffected.

When the prior distribution for \( \sigma_u^2 \) is \( \bar{\sigma}_u^2 \sim IG(\bar{\alpha}, \bar{\beta}) \), the posterior for \( \bar{\sigma}_u^2 \) is \( \bar{\sigma}_u^2 \sim IG(\hat{\alpha}, \hat{\beta}) \), where

\[
\hat{\alpha} = \bar{\alpha} + \frac{T}{2} \quad \text{and} \quad \hat{\beta} = \bar{\beta} + \frac{1}{2} \sum (\Delta p_t - \hat{c} \Delta q_t)^2 \tag{65} \]

and \( IG \) represents the inverted gamma distribution with parameters \( \alpha \) and \( \beta \).

The trade direction indicators \( q \) are drawn from the binomial distribution when \( c \) and \( \sigma_u \) are known. The first draw is \( q_1|q_2, \ldots, q_T \), the second draw is \( q_2|q_1, q_3, \ldots, q_T \), and the final draw is \( q_T|q_1, \ldots, q_{T-1} \). The full set of conditioning information includes the price changes \( \Delta p \) and the parameters \( c \) and \( \sigma_u \). Let \( f \) denote the normal density function with zero mean and volatility \( \sigma_u \). Then the posterior odds ratio of a buy versus a sell for the second trade
(q_0 \equiv 0 \text{ because the first trade of each day is reported as a midpoint price})\text{ is:}

\[
\frac{Pr(q_1 = +1|q_0, q_2, \ldots)}{Pr(q_1 = -1|q_0, q_2, \ldots)} = \frac{f(u_1(q_1 = +1))}{f(u_1(q_1 = -1))}.
\]

(66)

To draw \(q_2, \ldots, q_t, \ldots, q_{T-1}\), the following odds ratio is used:

\[
\frac{Pr(q_t = +1|q_1, q_2, \ldots)}{Pr(q_t = -1|q_1, q_2, \ldots)} = \frac{f(u_t(q_t = +1)) \cdot f(u_{t+1}(q_t = +1))}{f(u_t(q_t = -1)) \cdot f(u_{t+1}(q_t = +1))}.
\]

(67)

The final trade indicator of the day has the following odds ratio:

\[
\frac{Pr(q_T = +1|q_0, \ldots, q_{T-1})}{Pr(q_T = -1|q_0, \ldots, q_{T-1})} = \frac{f(w_T(q_T = +1))}{f(w_T(q_T = -1))}.
\]

(68)

**Specifics**

The estimation performed in this paper includes transactions data for each stock over a period of ten days. I simulate the Gibbs sampler separately for each stock. I initialize the priors to be: \((\hat{\mu}_c, \hat{\sigma}_c, \hat{\alpha}, \hat{\beta}) = (0, 0.005, 10^{-12}, 3.6 \cdot 10^{-15}/T)\), and draw \(c\) from a non-negative truncated normal distribution. This selection is associated with a mean bid-ask spread of approximately 0.8 percent and an annualized volatility of approximately 30 percent.

The beginning step initializes the sampler by selecting the starting values for \(q\). The first trade indicator of each day is set to 0 because the first quoted price is the price that clears all buys and sells at the market open. The remaining trade indicators are set to the sign of the most recent price change. Next \(\sigma_u^2\) is drawn from its prior distribution. The following steps are then repeated 1200 times; the first 200 iterations are discarded as “burn in”.

1. Using the most recent simulated value for \(\sigma_u^2\) and \(q\), the posterior of \(\hat{c}\) is computed using equation (64) and then \(c\) is drawn from its posterior distribution.

2. Using \(c\) and \(q\), the implied residuals \(u\) are calculated, the posterior of \(\hat{\sigma}_u^2\) is computed using equation (65), and then \(\sigma_u^2\) is drawn from its posterior distribution.

3. Using \(c\) and \(\sigma_u^2\), new draws for \(q\) are simulated. The first trade indicator for each day is set to 0. The second trade indicator for each day is drawn using the odds-
ratio provided by equation (66). All trade indicators except for the final indicator for each day are drawn using the odds-ratio provided by equation (67). The final trade indicator each day is drawn using the odds-ratio provided by equation (68).

The final estimate of the bid-ask spread for each stock is the mean of the 1000 draws of $c$ after the burn in is completed.
Penalized Method of Moments and Empirical Likelihood Estimation: Negative Weights and Shrinkage

Proof that weights may be implicitly defined as a function of the parameter vector. Let $F(\omega)$ be the $n$ dimensional vector of equations that defines the $n$ weights:

$$F_i(\omega) = \omega_i - \frac{1}{n} \left( 1 + \frac{\delta}{1-\delta} (g_i(\theta) - G_n(\theta))^t W_n G_n(\theta) \right) .$$

(69)

Then,

$$\frac{\partial F_i(\omega)}{\partial \omega_i} = 1 + \frac{\delta}{n(1-\delta)} \left( 1 + \frac{\delta}{1-\delta} (g_i(\theta) - G_n(\theta))^t W_n G_n(\theta) \right)^2 \times (g_i(\theta)^t W_n g_i(\theta) - 2g_i(\theta)^t W_n G_n(\theta))$$

(70)

$$\frac{\partial F_i(\omega)}{\partial \omega_j} = \frac{\delta}{n(1-\delta)} \left( 1 + \frac{\delta}{1-\delta} (g_i(\theta) - G_n(\theta))^t W_n G_n(\theta) \right)^2 \times (g_i(\theta)^t W_n g_j(\theta) - 2g_j(\theta)^t W_n G_n(\theta))$$

(71)

As $n \to \infty$, the matrix $\frac{\partial F(\omega)}{\partial \omega} \overset{p}{\to} I_n$, which is invertible. By the inverse function theorem, $\omega = \omega(\theta)$ is thus a continuous differentiable function of $\theta$.

Proof of Lemma 3.4.1. Equation (3.28) may be rewritten

$$n^{-1} = \omega_i + \omega_i \frac{\delta}{1-\delta} (g_i(\theta) - G_n(\theta))^t W_n G_n(\theta).$$

(72)
Multiply (72) by \( g_i(\theta) \) and sum over \( i \) to obtain

\[
G_n(\theta) = G_n(\theta) + \frac{\delta}{1-\delta} \left( \sum_i \omega_i g_i(\theta) g_i(\theta)' - G_n(\theta) G_n(\theta)' \right) W G_n(\theta)
\]

\[
= G_n(\theta) + \frac{\delta}{1-\delta} V_n(\theta) W_n G_n(\theta).
\]

Rearrange to obtain

\[
G_n(\theta) = (1 - \delta) (\delta V_n(\theta) W_n + (1 - \delta) I_m)^{-1} G_n(\theta)
\]

\[
= (1 - \delta) (\delta V_n(\theta) + (1 - \delta) W_n^{-1})^{-1} G_n(\theta)
\]

\[
= (1 - \delta) W_n^{-1} S_n(\theta) G_n(\theta),
\]

where the second line is due to the positive-definiteness of \( W_n \) and \( V_n(\theta) \).

**Proof of Lemma 3.4.3.** First, by the law of large numbers and central limit theorem \( V_n(\theta_0) \) and \( M_n(\theta_0) \) are consistent estimates of \( \Sigma \) and \( M \) respectively such that \( V_n(\theta_0) = \Sigma + O(n^{-\frac{1}{2}}) \) and \( M_n(\theta_0) = M + O(n^{-\frac{1}{2}}) \).

By equation (3.33)

\[
n(\omega_i(\theta_0) - n^{-1}) = -\frac{y}{1+y}
\]

where

\[
y = \delta g_i(\theta_0) S_n(\theta_0) G_n(\theta_0) - \delta (1 - \delta) G_n(\theta_0)' S_n(\theta_0) W_n^{-1} S_n(\theta_0) G_n(\theta_0).
\]

By Taylor expansion around \( y = 0 \):

\[
-\frac{y}{1+y} \approx -y - \frac{1}{2} y^2.
\]

Suppose \( V_n(\theta_0) = \Sigma + O(n^{-\frac{1}{2}}) \). Then, \( S_n(\theta_0) = \Sigma^{-1} + O(n^{-\frac{1}{2}}) \). Because \( G_n(\theta_0) = O(n^{-\frac{1}{2}}) \)

\[
\delta g_i(\theta_0) S_n(\theta_0) G_n(\theta_0) = \delta g_i(\theta_0) \left( \Sigma^{-1} + O\left(n^{-\frac{1}{2}}\right) \right) O\left(n^{-\frac{1}{2}}\right) = O\left(n^{-\frac{1}{2}}\right)
\]
Because the second term of $y$ is quadratic in $G_n(\theta_0)$, it can similarly be shown that

$$\delta(1 - \delta)G_n(\theta_0)\mathcal{S}_n(\theta_0)W_n^{-1}\mathcal{S}_n(\theta_0)G_n(\theta_0) = \mathcal{O}(n^{-1}).$$  \hfill (83)$$

Combining the two components of $y$, we have $n(\omega_i(\theta_0) - n^{-1}) = \mathcal{O}(n^{-\frac{1}{2}})$ if our initial assumption that $V_n(\theta_0) = \Sigma + \mathcal{O}(n^{-\frac{1}{2}})$ is true. By definition,

$$V_n(\theta_0) = \sum_i \omega_i(\theta_0)g_i(\theta_0)g_i(\theta_0)' - G_n(\theta_0)G_n(\theta_0)'$$  \hfill (84)$$

$$= V_n(\theta_0) + \sum_i (\omega_i(\theta_0) - n^{-1}) v_i(\theta_0) - (1 - \delta)^2G_n(\theta_0)\mathcal{S}_n(\theta_0)W_n^{-2}\mathcal{S}_n(\theta_0)G_n(\theta_0)$$

$$= \left(\Sigma + \mathcal{O}(n^{-\frac{1}{2}})\right) + \mathcal{O}\left(n^{-\frac{1}{2}}\right) + \mathcal{O}(n^{-1})$$  \hfill (85)$$

$$= \Sigma + \mathcal{O}\left(n^{-\frac{1}{2}}\right)$$  \hfill (86)$$

Hence, the assumed order of $V_n(\theta_0)$ is verified to be correct. Thus, part (iii) of the Lemma is proved directly, and the proof for part (i) is now complete. Further, we now have $\mathcal{S}_n(\theta_0) = \Sigma + \mathcal{O}(n^{-\frac{1}{2}})$. Part (ii) is proved as follows:

$$\sqrt{n}G_n(\theta_0) = (1 - \delta)W_n^{-1}\mathcal{S}_n(\theta_0)\sqrt{n}G_n(\theta_0)$$  \hfill (87)$$

$$= (1 - \delta)\left(\Sigma + \mathcal{O}\left(n^{-\frac{1}{2}}\right)\right)\left(\Sigma^{-1} + \mathcal{O}\left(n^{-\frac{1}{2}}\right)\right)\sqrt{n}G_n(\theta_0)$$  \hfill (88)$$

$$= (1 - \delta)\Sigma\Sigma^{-1}\sqrt{n}G_n(\theta_0) + \mathcal{O}\left(n^{-\frac{1}{2}}\right).$$  \hfill (89)$$

Simplifying the last equation in the series completes the proof. Finally, part (iv) of the Lemma is proved using part (i):

$$\mathcal{M}_n(\theta_0) = \sum_i \omega_i(\theta_0)m_i(\theta_0)$$  \hfill (90)$$

$$= \mathcal{M}_n(\theta_0) + \frac{1}{n}\sum_i n(\omega_i(\theta_0) - n^{-1}) m_i(\theta_0)$$  \hfill (91)$$

$$= \mathcal{M} + \mathcal{O}\left(n^{-\frac{1}{2}}\right) + \frac{1}{n}\sum_i \mathcal{O}\left(n^{-\frac{1}{2}}\right) m_i(\theta_0).$$  \hfill (92)$$

Simplifying the final equation in the series finishes the proof.
Proof of Theorem 3.4.4. The proof is similar to that provided by Qin and Lawless (1994). Denote \( \theta = \theta_0 + un^{-\varphi} \), for \( \theta \in \{ \theta \| \theta - \theta_0 \| = n^{-\varphi} \} \), where \( \| u \| = 1 \). The objective function may be rewritten

\[
Q_n(\theta) = n(1 - \delta)G_n(\theta)'S_n(\theta)W_n^{-1}S_n(\theta)G_n(\theta)
+ \frac{2}{\delta} \sum_i \ln \left( 1 + \delta g_i(\theta)'S_n(\theta)G_n(\theta) \right)
\]

(93)

Taylor approximating the log term by \( \ln(1 + x) = x - \frac{1}{2}x^2 + o(x^2) \), the objective function is approximated by

\[
Q_n(\theta) = n(1 - \delta)G_n(\theta)'S_n(\theta)W_n^{-1}S_n(\theta)G_n(\theta) + 2nG_n(\theta)'S_n(\theta)G_n(\theta)
- 2n(1 - \delta)G_n(\theta)'S_n(\theta)W_n^{-1}S_n(\theta)G_n(\theta)
- n\delta G_n(\theta)'S_n(\theta)V_n(\theta)S_n(\theta)G_n(\theta)
+ o(nG_n(\theta)'S_n(\theta)W_n^{-1}S_n(\theta)G_n(\theta))
\]

(94)

Using the definition of \( S_n(\theta) \), the following expression emerges:

\[
Q_n(\theta) = nG_n(\theta)'S_n(\theta)G_n(\theta) + o(nG_n(\theta)'S_n(\theta)W_n^{-1}S_n(\theta)G_n(\theta)).
\]

(95)

By Taylor expansion around \( \theta_0 \), we have (uniformly for \( u \)),

\[
Q_n(\theta) = n \left[ G_n(\theta_0) + M_n(\theta_0)un^{-\varphi} \right]' S_n(\theta) \left[ G_n(\theta_0) + M_n(\theta_0)un^{-\varphi} \right]
+ o \left( n \left[ G_n(\theta_0) + M_n(\theta_0)un^{-\varphi} \right]' S_n(\theta)W_n^{-1}S_n(\theta) \left[ G_n(\theta_0) + M_n(\theta_0)un^{-\varphi} \right] \right).
\]

(96)

By the law of iterated logarithms and \( \varphi < \frac{1}{2} \),

\[
Q_n(\theta) = n \left[ \mathcal{O}(n^{-\frac{1}{2}}(\log \log n)^{\frac{1}{2}}) + E[M_n(\theta_0)]un^{-\varphi} \right]' \times \left[ \delta E[v_i(\theta_0)] + (1 - \delta)W_n^{-1} \right]^{-1}
\times \left[ \mathcal{O}(n^{-\frac{1}{2}}(\log \log n)^{\frac{1}{2}}) + E[M_n(\theta_0)]un^{-\varphi} \right] + o \left( n^{1-2\varphi} \right)
\geq (c - \varepsilon)n^{1-2\varphi},
\]

(97)
where \( c - \varepsilon \geq 0 \) and \( c \) is the smallest eigenvalue of

\[
E \left( m_i(\theta_0) \right)' \left[ \delta E[v_i(\theta_0)] + (1 - \delta)W_n^{-1} \right]^{-1} E \left( m_i(\theta_0) \right).
\]  

(98)

Substituting \( \theta_0 \) into equation (95),

\[
Q_n(\theta_0) = nG_n(\theta_0)'S_n(\theta_0)G_n(\theta_0) + o(nG_n(\theta_0)'S_n(\theta_0)W_n^{-1}S_n(\theta_0)G_n(\theta_0)) \]

(99)

\[
= O(\log \log n) + o(1)
\]

(100)

\[
= O(\log \log n).
\]

(101)

Because the objective function is continuous around \( \theta \) as \( \theta \) belongs to the ball \( \| \theta - \theta_0 \| \leq n^{-\phi} \) and on the surface of the ball the objective function is order \( O(n^{1-2\phi}) \) while the order of the objective function at the population parameter value is \( O(\log \log n) \), the objective function achieves its minimum value within the interior of the ball.

**Proof of Theorem 3.4.5.** The proof follows Newey and McFadden (1994)'s proof of asymptotic normality for GMM. By assumptions (1) and (2), with probability approaching one the first-order conditions \( M_n(\hat{\theta})'W_nG_n(\hat{\theta}) = 0 \) are satisfied. Expand \( G_n(\theta) \) around \( \theta_0 \) to obtain

\[
G_n(\hat{\theta}) = G_n(\theta_0) + M_n(\bar{\theta})' (\hat{\theta} - \theta_0),
\]

(102)

where \( \bar{\theta} \) represents a mean value. Substitute in the relationship

\[
G_n(\hat{\theta}) = \frac{1}{1 - \delta}S_n(\hat{\theta})^{-1}W_nG_n(\hat{\theta})
\]

(103)

and multiply by \( M_n(\hat{\theta})'S_n(\hat{\theta}) \) to obtain

\[
0 = \frac{1}{1 - \delta}M_n(\hat{\theta})'W_nG_n(\hat{\theta}) = M_n(\hat{\theta})'S_n(\bar{\theta})G_n(\theta_0) + M_n(\hat{\theta})'S_n(\bar{\theta})M_n(\bar{\theta})(\hat{\theta} - \theta_0).
\]

(104)

Rearrange to obtain

\[
\sqrt{n}(\hat{\theta} - \theta_0) = \left( M_n(\hat{\theta})'S_n(\bar{\theta})M_n(\bar{\theta}) \right)^{-1}M_n(\hat{\theta})'S_n(\bar{\theta})G_n(\theta_0)
\]

(105)
By assumption (4) and \( \hat{\theta} \xrightarrow{p} \theta_0 \), with probability approaching one

\[
\| M_n(\hat{\theta}) - M \| \leq \| M_n(\hat{\theta}) - M(\hat{\theta}) \| + \| M(\hat{\theta}) - M \| \quad (106)
\]

\[
\leq \sup_{\theta \in \Theta} \| M_n(\theta) - M(\theta) \| + \| M(\hat{\theta}) - M \| \quad (107)
\]

\[
\xrightarrow{p} 0. \quad (108)
\]

Similarly, by assumption (4), \( \hat{\theta} \xrightarrow{p} \theta_0 \), and Lemma 3.4.3 (iv) with probability approaching one

\[
\| M_n(\hat{\theta}) - M \| \leq \| M_n(\hat{\theta}) - M(\hat{\theta}) \| + \| M(\hat{\theta}) - M \| + \| M - M \| \quad (109)
\]

\[
\leq \sup_{\theta \in \Theta} \| M_n(\theta) - M(\theta) \| + \| M(\hat{\theta}) - M^* \| + \| M - M \| \quad (110)
\]

\[
\xrightarrow{p} 0. \quad (111)
\]

Applying (3), (6), and the Slutsky theorem yields the result.

**Proof of Theorem 3.4.6.** Hansen and Newey and McFadden show that when the asymptotic variance of \( \sqrt{n}(\hat{\theta} - \theta_0) \) takes the form \( ((M' SM)^{-1} M' S \Sigma S M (M' SM)^{-1}) \), the minimum variance is obtained when \( S = \Sigma^{-1} \). Because \( W_n \xrightarrow{p} \Sigma^{-1} \) by assumption and \( \mathcal{V}_n(\theta_0) \xrightarrow{p} \Sigma \), we verify that \( S_n(\theta_0) \xrightarrow{p} \Sigma^{-1} \) the asymptotic variance reduces to \( (M' SM)^{-1} \).

**Proof of Theorem 3.4.7.** PMM’s objective function has two components, the quadratic penalty and the KLIC penalty.

By Lemma 3.4.1, the quadratic component of the objective function is rewritten:

\[
n(1 - \delta)G_n(\hat{\theta})' S_n(\hat{\theta}) W_n^{-1} S_n(\hat{\theta}) G_n(\hat{\theta}). \quad (112)
\]

Taylor approximating logarithms by \( ln(1 + x) = x - \frac{1}{2} x^2 + \mathcal{O}(x^3) \) and substituting in the optimal PMM weights conditioned on \( \theta \) given by equation (3.33), I obtain:

\[
\ln(n\omega_i(\hat{\theta})) = \delta g_i(\theta)' S_n(\hat{\theta}) G_n(\theta) - \delta(1 - \delta)G_n(\hat{\theta}) S_n(\hat{\theta}) W_n^{-1} S_n(\hat{\theta}) G_n(\theta) - \frac{1}{2} \delta^2 G_n(\hat{\theta}) S_n(\hat{\theta}) g_i(\hat{\theta}) g_i(\hat{\theta})' S_n(\hat{\theta}) G_n(\hat{\theta}) + \mathcal{O}(n^{-\beta}). \quad (113)
\]
Summing equation (113) over $i$ provides

$$
\sum_i \ln(n\omega_i(\hat{\theta})) = \delta nG_n(\hat{\eta})G_n(\theta) - n\delta(1 - \delta)G_n(\hat{\theta})S_n(\hat{\theta})W_n^{-1}S_n(\hat{\theta})G_n(\theta) \\
- \frac{n}{2}\delta^2G_n(\hat{\theta})S_n(\hat{\theta})V_n(\hat{\theta})S_n(\hat{\theta})G_n(\hat{\theta}) + O(n^{-\frac{3}{2}}).
$$

(114)

Dividing equation (112) by $1 - \delta$ and equation (114) by $-\delta^2$, adding the two components, and simplifying the expression,

$$
Q(\hat{\theta})_{\text{PMM}} = nG_n(\hat{\theta})'S_n(\hat{\theta})G_n(\hat{\theta}) + O \left( n^{-\frac{3}{2}} \right). 
$$

(115)

The quadratic statistic is singular:

$$
nG_n(\hat{\theta})'S_n(\hat{\theta})G_n(\hat{\theta}) = nG_n(\hat{\theta})'S_n(\hat{\theta})^{\frac{1}{2}}M_n(\hat{\theta})^{\frac{1}{2}}S_n(\hat{\theta})^{\frac{1}{2}}G_n(\hat{\theta}) \\
= nG_n(\hat{\theta})'S_n(\hat{\theta})^{\frac{1}{2}} \left( P_{\mathcal{M}(\hat{\eta})} + P_{\mathcal{M}(\hat{\eta})}^\perp \right) S_n(\hat{\theta})^{\frac{1}{2}}G_n(\hat{\theta}) \\
= nG_n(\hat{\theta})'S_n(\hat{\theta})^{\frac{1}{2}}P_{\mathcal{M}(\hat{\eta})}S_n(\hat{\theta})^{\frac{1}{2}}G_n(\hat{\theta}) \\
= \text{rank}(m - k) 
$$

(116)

where $P_{\mathcal{M}(\hat{\eta})} = S_n(\hat{\theta})^{\frac{1}{2}}M_n(\hat{\theta}) \left( M_n(\hat{\eta})'S_n(\hat{\theta})M_n(\hat{\theta}) \right)^{-1}M_n(\hat{\eta})'S_n(\hat{\theta})^{\frac{1}{2}}$ is the projection matrix that sets $k$ dimensions of $G_n(\hat{\theta})$ to zero and $S_n(\hat{\theta})^{\frac{1}{2}}$ is the Cholesky decomposition of $S_n(\hat{\theta})$. If $W_n \overset{p}{\rightarrow} \Sigma^{-1}$ then $S_n(\hat{\theta}) \overset{p}{\rightarrow} \Sigma^{-1}$.

**Proof of Theorem 3.4.8.** The final form of the higher order asymptotic expansion is the result of a number of linearizations. First, I provide linear approximations of $S_n(\hat{\theta})$ and $M_n(\hat{\theta})$, the derivations of which are included at the end of proof.

$$
S_n(\hat{\theta}) = \Sigma^{-1} - \Sigma^{-1} \left( V_n(\hat{\theta}) - \Sigma \right) \Sigma^{-1} \\
+ \Sigma^{-1} \left( (\delta)^2 \frac{1}{n} \sum_i g_i(\hat{\theta})' \Sigma^{-1} G_n(\hat{\theta}) v_i(\hat{\theta}) + (1 - \delta) \sum_j \hat{\Gamma}_j G_n(\hat{\theta})' (\tilde{\Upsilon} - \Upsilon)' e_j \right) \Sigma^{-1} \\
+ O(n^{-1})
$$

(117)
and

\[
\mathcal{M}_n(\hat{\theta}) = M + \left( M_n(\hat{\theta}) - \frac{1}{n} \sum_i m_i(\hat{\theta}) g_i(\hat{\theta}) \Sigma^{-1} G_n(\hat{\theta}) - M \right) + O(n^{-1}) ,
\]

(118)

where \( \hat{\Gamma}_j = \frac{\partial V_n(\hat{\theta})}{\partial \theta^j} \). Note that in each equation, the first term is the limiting value of the respective equation as \( n \to \infty \) evaluated at the true parameter value and is \( O(1) \), the final term is \( O(n^{-1}) \), and the middle term(s) is the estimation error and is \( O(n^{-1/2}) \). The sample average Jacobian matrix is rewritten as

\[
M_n(\hat{\theta}) = M + \left( M_n(\hat{\theta}) - M \right).
\]

(119)

Again the first term is the limiting value as \( n \to \infty \) of the estimated Jacobian evaluated at the true parameter value, the second term is the estimation error, which is \( O(n^{-1/2}) \), and in this case there is no approximation error. The next step is to expand the sample average moment condition \( G_n(\hat{\theta}) \) around the true value \( \theta_0 \):

\[
G_n(\hat{\theta}) = G_n(\theta_0) + M_n(\hat{\theta}) \left( \hat{\theta} - \theta_0 \right) + 0.5H_n(\hat{\theta}) \left[ \left( \hat{\theta} - \theta_0 \right) \otimes \left( \hat{\theta} - \theta_0 \right) \right] + O\left(n^{-\frac{3}{2}}\right),
\]

(120)

Equation (3.29) provides the first-order condition that \( \mathcal{M}_n(\hat{\theta})^\prime S_n(\hat{\theta}) G_n(\hat{\theta}) = 0 \). Multiply (120) by \( \mathcal{M}_n(\hat{\theta})^\prime S_n(\hat{\theta}) \) to eliminate the left hand side:

\[
0 = \mathcal{M}_n(\hat{\theta})^\prime S_n(\hat{\theta}) G_n(\theta_0) + \mathcal{M}_n(\hat{\theta})^\prime S_n(\hat{\theta}) M_n(\hat{\theta}) \left( \hat{\theta} - \theta_0 \right)
\]

\[
+ 0.5 \mathcal{M}_n(\hat{\theta})^\prime S_n(\hat{\theta}) H_n(\hat{\theta}) \left[ \left( \hat{\theta} - \theta_0 \right) \otimes \left( \hat{\theta} - \theta_0 \right) \right] + O\left(n^{-\frac{3}{2}}\right)
\]

(121)

Define

\[
\Omega_n(\hat{\theta}) \equiv \mathcal{M}_n(\hat{\theta})^\prime S_n(\hat{\theta}) M_n(\hat{\theta}) \quad \quad \Omega_n^{-1}(\hat{\theta}) \mathcal{M}_n(\hat{\theta})^\prime S_n(\hat{\theta})
\]

(122)

and rearrange equation (121) to obtain

\[
\hat{\theta} - \theta_0 = -\Upsilon_n(\hat{\theta}) G_n(\theta_0) - 0.5 \Upsilon_n(\hat{\theta}) H_n(\hat{\theta}) \left[ \left( \hat{\theta} - \theta_0 \right) \otimes \left( \hat{\theta} - \theta_0 \right) \right] + O\left(n^{-\frac{3}{2}}\right)
\]

(123)
Next, approximate \( \Omega_n^{-1}(\hat{\theta}) \) with a two-term Taylor Expansion:

\[
\begin{align*}
\Omega_n^{-1}(\hat{\theta}) &= \Omega^{-1} - \Omega^{-1} \left( \Omega_n(\hat{\theta}) - \Omega \right) \Omega^{-1} + \mathcal{O} \left( n^{-1} \right) \\
&= \Omega^{-1} - \Omega^{-1} \left( M_n(\hat{\theta}) - M \right)^{'} \Sigma^{-1} M + M' \left( S_n(\hat{\theta}) - \Sigma^{-1} \right) M \Omega^{-1} \\
&\quad - \Omega^{-1} \left( M' \Sigma^{-1} \left( M_n(\hat{\theta}) - M \right) \right) \Omega^{-1} + \mathcal{O} \left( n^{-1} \right) \\
&= \Omega^{-1} - \Omega^{-1} \left( M_n(\hat{\theta}) - \frac{1}{n} \sum_i m_i(\hat{\theta}) g_i(\hat{\theta})^{'} \Sigma^{-1} G_n(\hat{\theta}) - M \right)^{'} \Sigma^{-1} M \Omega^{-1} \\
&\quad + \Upsilon \left( V_n(\hat{\theta}) - \Sigma \right) \Sigma^{-1} M \Omega^{-1} \\
&\quad - \Upsilon \left( \frac{\delta^2}{n} \sum_i g_i(\hat{\theta})^{'} \Sigma^{-1} G_n(\hat{\theta}) v_i(\hat{\theta}) + (1 - \delta) \sum_{j=1}^k \hat{\Gamma}_j G_n(\theta_0)^{'} (\hat{\Upsilon} - \Upsilon)^{'} e_j \right) \Sigma^{-1} M \Omega^{-1} \\
&\quad - \Upsilon \left( M_n(\hat{\theta}) - M \right) \Omega^{-1} + \mathcal{O} \left( n^{-1} \right).
\end{align*}
\]

The final equation in the series is the result of substituting in the approximations provided by equations (143) to (119). The first term is \( \mathcal{O}(1) \) and the next three terms are \( \mathcal{O} \left( n^{-1/2} \right) \).

In equation (123), the first term includes \( G_n(\theta_0) \) which is \( \mathcal{O} \left( n^{-1/2} \right) \) and the second term includes \( (\hat{\theta} - \theta_0) \otimes (\hat{\theta} - \theta_0) \) which is \( \mathcal{O} \left( n^{-1} \right) \). Hence, because we are investigating the \( \mathcal{O} \left( n^{-1} \right) \) properties of the estimator \( \hat{\theta} \), for the first term, we only require the \( \mathcal{O} \left( n^{-1/2} \right) \) components of \( \Omega_n^{-1}(\hat{\theta}) \), \( M_n(\hat{\theta}) \), and \( S_n(\hat{\theta}) \). For the second term, only their limiting values as \( n \to \infty \) evaluated at the true parameter value \( \theta_0 \) are needed. The rest of the terms are
\[ O \left( n^{-3/2} \right) \). After substituting in the relevant terms, I obtain

\[
\hat{\theta} - \theta_0 = - \Upsilon G_n(\theta_0)
\]

\[
+ \Omega^{-1} \left( M_n(\hat{\theta}) - \delta \frac{1}{n} \sum_i m_i(\hat{\theta}) g_i(\hat{\theta})' \Sigma^{-1} G_n(\hat{\theta}) - M \right)' \Sigma^{-1} M \Omega^{-1} M' \Sigma^{-1} G_n(\theta_0)
\]

\[
- \Upsilon \left( V_n(\hat{\theta}) - \delta \frac{1}{n} \sum_i g_i(\hat{\theta})' \Sigma^{-1} G_n(\hat{\theta}) v_i(\hat{\theta}) - \Sigma \right) \Sigma^{-1} M \Omega^{-1} M' \Sigma^{-1} G_n(\theta_0)
\]

\[
+ \Upsilon \left( (1 - \delta) \sum_{j=1}^k \hat{\Gamma}_j G_n(\theta_0)' (\hat{\Upsilon} - \Upsilon)' e_j \right) \Sigma^{-1} M \Omega^{-1} M' \Sigma^{-1} G_n(\theta_0)
\]

\[
+ \Upsilon \left( M_n(\hat{\theta}) - M \right) \Upsilon G_n(\theta_0)
\]

\[
- \Omega^{-1} \left( M_n(\hat{\theta}) - \delta \frac{1}{n} \sum_i m_i(\hat{\theta}) g_i(\hat{\theta})' \Sigma^{-1} G_n(\hat{\theta}) - M \right) \Sigma^{-1} G_n(\theta_0)
\]

\[
+ \Upsilon \left( V_n(\hat{\theta}) - \delta \frac{1}{n} \sum_i g_i(\hat{\theta})' \Sigma^{-1} G_n(\hat{\theta}) v_i(\hat{\theta}) - \Sigma \right) \Sigma^{-1} G_n(\theta_0)
\]

\[
- \Upsilon \left( (1 - \delta) \sum_{j=1}^k \hat{\Gamma}_j G_n(\theta_0)' (\hat{\Upsilon} - \Upsilon)' e_j \right) \Sigma^{-1} G_n(\theta_0)
\]

\[
- 0.5 \Upsilon H_n(\hat{\theta}) \left[ (\hat{\theta} - \theta_0) \otimes (\hat{\theta} - \theta_0) \right]
\]

\[
+ O \left( n^{-1} \right).
\]

The first line of (125) is \( O \left( n^{-1/2} \right) \) and lines two through nine are \( O \left( n^{-1} \right) \). Lines two through five are from the \( O \left( n^{-1/2} \right) \) component of the expansion of \( \Omega_n^{-1}(\hat{\theta}) \). The second line is due to \( M_n(\hat{\theta})' \). Lines three and four are attributed to \( S_n(\hat{\theta}) \). The fifth line is from \( M_n(\hat{\theta}) \). In addition, \( O \left( n^{-1/2} \right) \) terms are embedded in the \( M_n(\hat{\theta})' \) and \( S_n(\hat{\theta}) \), the other two components of \( \Upsilon_n(\hat{\theta}) \). Line six is due to \( M_n(\hat{\theta})' \) and lines seven and eight are from \( S_n(\hat{\theta}) \). Finally, the ninth line is the third term in the expansion of \( G_n(\hat{\theta}) \). In order to calculate the
$\mathcal{O}(n^{-1})$ bias, take expectations of (125) and rearrange to obtain:

$$
\mathbb{E} \left[ \hat{\theta} - \theta_0 \right] = \mathbb{E} \left[ M_n(\hat{\theta}) \gamma G_n(\theta_0) \right] - 0.5 \mathbb{Y} \mathbb{H} \mathbb{E} \left[ \left( \hat{\theta} - \theta_0 \right) \otimes \left( \hat{\theta} - \theta_0 \right) \right] \\
- \Omega^{-1} \mathbb{E} \left[ \left( M_n(\hat{\theta}) - \delta \frac{1}{n} \sum_i m_i(\hat{\theta}) g_i(\hat{\theta})' \Sigma^{-1} G_n(\hat{\theta}) \right)' \Sigma^{-\frac{1}{2}} P_{\frac{1}{2} M} \Sigma^{-\frac{1}{2}} G_n(\theta_0) \right] \\
+ \mathbb{Y} \mathbb{E} \left[ \left( V_n(\hat{\theta}) - \delta^2 \frac{1}{n} \sum_i g_i(\hat{\theta})' \Sigma^{-1} G_n(\hat{\theta}) v_i(\hat{\theta}) \right) \Sigma^{-\frac{1}{2}} P_{\frac{1}{2} M} \Sigma^{-\frac{1}{2}} G_n(\theta_0) \right] \\
- \mathbb{Y} \mathbb{E} \left[ \left( 1 - \delta \right) \sum_{j=1}^k \Gamma_j G_n(\theta_0)' (\hat{\Upsilon} - \Upsilon)' e_j \right] \Sigma^{-\frac{1}{2}} P_{\frac{1}{2} M} \Sigma^{-\frac{1}{2}} G_n(\theta_0) \\
+ \mathcal{O} \left( n^{-\frac{3}{4}} \right),
$$

where $P_{\frac{1}{2} M} \equiv \mathbb{I}_m - \Sigma^{-\frac{1}{2}} M (M' \Sigma^{-1} M)^{-1} \Sigma^{-\frac{1}{2}} M$ is a projection matrix orthogonal to the space spanned by the asymptotic normalized Jacobian. Before proceeding, I note:

$$
g_i(\hat{\theta}) = g_i(\theta_0) + \mathcal{O} \left( n^{-\frac{3}{4}} \right) \tag{127}
$$

$$
m_i(\hat{\theta}) = m_i(\theta_0) + \mathcal{O} \left( n^{-\frac{3}{4}} \right) \tag{128}
$$

$$
G_n(\hat{\theta}) = G_n(\theta_0) + M' (\hat{\theta} - \theta_0) + \left( M_n(\hat{\theta}) - M \right)' \left( \hat{\theta} - \theta_0 \right) + \mathcal{O} \left( n^{-1} \right) \tag{129}
$$

$$
nG_n(\theta_0) G_n(\theta_0)' \Sigma^{-1} = \mathbb{I}_m + \left( nG_n(\theta_0) G_n(\theta_0)' \Sigma^{-1} - \mathbb{I}_m \right) / \mathcal{O}(n^{-1/2}) \tag{130}
$$

Using the result of the expansion of $G_n(\hat{\theta})$: $\hat{\theta} - \theta_0 = -(M' \Sigma^{-1} M)^{-1} M' \Sigma^{-1} G_n(\theta_0) + \mathcal{O} \left( n^{-1} \right)$, the assumption that the observations are independent, and equation (128), the first line of (126) may be rewritten

$$
B_I = n^{-1} \mathbb{Y} \left( \mathbb{E} \left[ m_i(\theta_0) \gamma g_i(\theta_0) \right] - a \right), \quad \tag{131}
$$

where $a$ is an $m \times 1$ matrix such that

$$
a_j \equiv 0.5 \text{tr} \left( \Omega^{-1} \mathbb{E} \left[ \partial^2 g_{ij}(\theta_0) / \partial \theta \partial \theta' \right] \right) \quad (j = 1, \ldots, m), \tag{132}
$$

and $g_{ij}(\theta_0)$ represents the $j$th element of $g_i(\theta)$. 

157
The next step is to rewrite the second line of (126). The three approximations given by equations (127) through (129) are substituted into the second line of (126). Note that the second term of (128) is orthogonal to $P_{\frac{n}{M}}$ and the third term is $O\left(n^{-1}\right)$.

$$-\Omega^{-1}E \left[ \left( M_n(\hat{\theta}) - \delta \frac{1}{n} \sum_i m_i(\hat{\theta}) g_i(\theta)'\Sigma^{-1}G_n(\theta) \right) \Sigma^{-\frac{1}{2}}P_{\frac{n}{M}}\Sigma^{-\frac{1}{2}}G_n(\theta_0) \right]$$

$$= -\Omega^{-1}E \left[ \frac{1}{n} \sum_i m_i(\hat{\theta})'\Sigma^{-\frac{1}{2}}P_{\frac{n}{M}}\Sigma^{-\frac{1}{2}}G_n(\theta_0) \right]$$

$$+ \Omega^{-1}E \left[ \frac{1}{n} \sum_i m_i(\hat{\theta})'\Sigma^{-\frac{1}{2}}P_{\frac{n}{M}}\Sigma^{-\frac{1}{2}}G_n(\theta_0)G_n(\theta_0)'\Sigma^{-\frac{1}{2}}P_{\frac{n}{M}}\Sigma^{-\frac{1}{2}}G_n(\theta_0) \right] + O\left(n^{-\frac{3}{2}}\right)$$

$$= -\Omega^{-1}E \left[ \frac{1}{n} \sum_i m_i(\hat{\theta})'\Sigma^{-\frac{1}{2}}P_{\frac{n}{M}}\Sigma^{-\frac{1}{2}}g_i(\theta_0) \right]$$

$$+ \Omega^{-1}E \left[ \frac{1}{n} \sum_i m_i(\hat{\theta})'\Sigma^{-\frac{1}{2}}P_{\frac{n}{M}}\Sigma^{-\frac{1}{2}}g_i(\theta_0) \right] + O\left(n^{-\frac{3}{2}}\right)$$

$$= -\Omega^{-1}E \left[ \frac{1}{n} \sum_i m_i(\hat{\theta})'\Sigma^{-\frac{1}{2}}P_{\frac{n}{M}} \left( \Sigma_m - \delta \Sigma^{-\frac{1}{2}} \left( nG_n(\theta_0)G_n(\theta_0)' \right) \Sigma^{-\frac{1}{2}} \right) \Sigma^{-\frac{1}{2}}g_i(\theta_0) \right]$$

$$+ O\left(n^{-\frac{3}{2}}\right)$$

$$= -\Omega^{-1}E \left[ \frac{1}{n^2} \sum_i m_i(\hat{\theta})'\Sigma^{-\frac{1}{2}}P_{\frac{n}{M}} \left( \Sigma_m - \delta \Sigma^{-\frac{1}{2}}g_i(\theta_0) \right) \right] + O\left(n^{-\frac{3}{2}}\right)$$

$$= -n^{-1} (1 - \delta) \Omega^{-1}E \left[ m_i(\hat{\theta})'\Sigma^{-\frac{1}{2}}P_{\frac{n}{M}}\Sigma^{-\frac{1}{2}}g_i(\theta_0) \right] + O\left(n^{-\frac{3}{2}}\right)$$

$$= -n^{-1} (1 - \delta) \Omega^{-1}E \left[ m_i(\hat{\theta})'\Sigma^{-\frac{1}{2}}P_{\frac{n}{M}}\Sigma^{-\frac{1}{2}}g_i(\theta_0) \right] + O\left(n^{-\frac{3}{2}}\right)$$

The second line substitutes in approximations (127) and (129). The third line distributes $\Sigma^{-1/2}P_{\frac{n}{M}}\Sigma^{-1/2}G_n(\theta_0)$. The fourth line moves and transposes the scalar value $g_i(\theta_0)'\Sigma^{-1}G_n(\theta_0)$ within the second summation. The fifth line simplifies the left summation using the independence assumption and multiplies and divides the second summation by $n$. The sixth line simplifies and the seventh line substitutes in the approximation given by (130). The last line provides the result after substituting in approximation (128). Note that my notation
differs from Newey and Smith (2004), who denote their result by $B_G$.

The exact same procedure for the third line of equation (126) provides the bias term associated with the estimation of the second moment matrix

$$B_\Sigma = n^{-1} (1 - \delta^2) \mathcal{Y} \mathbb{E} \left[ v_i(\theta_0) \Sigma^{-\frac{1}{2}} \mathbf{P}_M \Sigma^{-\frac{1}{2}} g_i(\theta_0) \right]$$  \hspace{1cm} (133)

Finally, I rewrite the fourth line of equation (126):

$$B_W = -\mathcal{Y} \mathbb{E} \left[ (1 - \delta) \sum_{j=1}^{k} \tilde{\Gamma}_j G_n(\theta_0)'(\bar{\mathcal{Y}} - \mathcal{Y})' e_j \right] \Sigma^{-\frac{1}{2}} \mathbf{P}_M \Sigma^{-\frac{1}{2}} G_n(\theta_0)$$

$$= -\mathcal{Y} \mathbb{E} \left[ (1 - \delta) \sum_{j=1}^{k} \tilde{\Gamma}_j \Sigma^{-\frac{1}{2}} \mathbf{P}_M \Sigma^{-\frac{1}{2}} G_n(\theta_0) G_n(\theta_0)'(\bar{\mathcal{Y}} - \mathcal{Y})' e_j \right]$$

$$= -\mathcal{Y} \mathbb{E} \left[ (1 - \delta) \sum_{j=1}^{k} \tilde{\Gamma}_j \mathbf{P}_M \Sigma^{-1} (n G_n(\theta_0) G_n(\theta_0)')(\bar{\mathcal{Y}} - \mathcal{Y})' e_j \right]$$

$$= -\mathcal{Y} \mathbb{E} \left[ (1 - \delta) \sum_{j=1}^{k} \tilde{\Gamma}_j \mathbf{P}_M \Sigma^{-1} (n G_n(\theta_0) G_n(\theta_0)')(\bar{\mathcal{Y}} - \mathcal{Y})' e_j \right] + \mathcal{O} \left( n^{-\frac{3}{2}} \right)$$

$$= -\mathcal{Y} \mathbb{E} \left[ \sum_{j=1}^{k} \tilde{\Gamma}_j \mathbf{P}_M (\bar{\mathcal{Y}} - \mathcal{Y})' e_j \right] + \mathcal{O} \left( n^{-\frac{3}{2}} \right)$$

$$= -\mathcal{Y} \mathbb{E} \left[ \sum_{j=1}^{k} \tilde{\Gamma}_j (\bar{\mathcal{Y}} - \mathcal{Y})' e_j \right] + \mathcal{O} \left( n^{-\frac{3}{2}} \right)$$  \hspace{1cm} (134)

The second line moves the scalar value $G_n(\theta_0)'(\bar{\mathcal{Y}} - \mathcal{Y})' e_j$. The third line multiplies and divides by $n$. The fourth line is possible because both $\Sigma^{-1}$ and $\mathbf{P}_M$ are symmetric. The
fifth lines substitutes in the approximation provided by equation (130). The final result is:

\[
\text{Bias} \left( \hat{\theta}_{\text{PMM}} \right) = B_I + B_M + B_\Sigma + B_W
\]

\[
B_I = n^{-1} \mathbb{E} \left[ m_i(\theta_0) Y g_i(\theta_0) \right] - a
\]

\[
B_M = -n^{-1} (1 - \delta) \Omega^{-1} \mathbb{E} \left[ m_i(\theta_0) \Sigma^{-\frac{1}{2}} P_M \Sigma^{-\frac{1}{2}} g_i(\theta_0) \right]
\]

\[
B_\Sigma = n^{-1} (1 - \delta^2) \mathbb{E} \left[ v_i(\theta_0) \Sigma^{-\frac{1}{2}} P_M \Sigma^{-\frac{1}{2}} g_i(\theta_0) \right]
\]

\[
B_W = -n^{-1} (1 - \delta) \mathbb{E} \left[ \sum_{j=1}^k \hat{\Gamma}_j \right] (\hat{\mathbf{Y}} - \mathbf{Y})^t e_j
\]

The Three Linearizations: \( n \omega_i(\hat{\theta}), \mathcal{S}_n(\hat{\theta}), \) and \( \mathcal{M}_n(\hat{\theta}) \)

The first step in the two linearizations is to linearize the probability weights \( \omega_i(\hat{\theta}) \). Recall the definition of the optimal PMM weights conditional on \( \theta \)

\[
\omega_i(\theta) = \frac{1}{n} \frac{1 + \delta g_i(\theta)^t \mathcal{S}_n(\theta) G_n(\theta) - n^{-1} \mathcal{S}_n(\theta) W_n^{-1} \mathcal{S}_n(\theta) G_n(\theta)}{(1 + \delta)^2 \mathcal{S}_n(\theta) W_n^{-1} \mathcal{S}_n(\theta) G_n(\theta)} + \mathcal{O}(n^{-1/2})
\]

A geometric expansion provides the linearized weights:

\[
\omega_i(\theta) = \frac{1}{n} \frac{1 - \delta g_i(\theta)^t \mathcal{S}_n(\theta) G_n(\theta)}{(1 + \delta)^2 \mathcal{S}_n(\theta) W_n^{-1} \mathcal{S}_n(\theta) G_n(\theta)} + \mathcal{O}(n^{-2}).
\]

Equation (137) can be further refined because \( \mathcal{S}_n(\theta) \) is an estimate of the second moment matrix. Recall the definition of \( \mathcal{S}_n(\theta) \)

\[
\mathcal{S}_n(\theta) = \left( \delta \mathcal{V}_n(\hat{\theta}) + (1 - \delta) W_n^{-1} \right)^{-1}.
\]

Because \( W_n^{-1} \) is a consistent estimate of \( \Sigma \), \( \mathcal{S}_n(\hat{\theta}) \) may be approximated by

\[
\mathcal{S}_n(\theta) = \Sigma^{-1} - \Sigma^{-1} \left( \delta \mathcal{V}_n(\hat{\theta}) + (1 - \delta) W_n^{-1} - \Sigma \right) \Sigma^{-1} + \mathcal{O}(n^{-1/2}).
\]

Before continuing with the approximation \( \mathcal{S}_n(\hat{\theta}) \), I substitute equation (139) into (??) to
obtain
\[
\omega_i(\hat{\theta}) = \frac{1}{n} - \frac{1}{n} \delta g_i(\hat{\theta})' \Sigma^{-1} G_n(\hat{\theta}) + \mathcal{O}(n^{-2}) .
\] (140)

Next, use equation (140) to approximate \( V_n(\hat{\theta}) \)
\[
V_n(\hat{\theta}) = \sum_i \omega_i(\hat{\theta}) v_i(\hat{\theta})
= V_n(\hat{\theta}) - \frac{1}{n} \sum_i v_i(\hat{\theta}) g_i(\hat{\theta})' \Sigma^{-1} G_n(\hat{\theta}) + \mathcal{O}(n^{-1})
\] (141)

and rewrite \( W_n^{-1} = V_n(\hat{\theta}) \), where \( \hat{\theta} \) is a first-round estimator of \( \theta \), as
\[
W_n^{-1} = V_n(\hat{\theta}) + \left( V_n(\hat{\theta}) - V_n(\hat{\theta}) \right)
= V_n(\hat{\theta}) - \sum_j^k \frac{\partial V_n(\theta_0)}{\partial \theta_j} \left( \hat{\theta}_j - \theta_{0j} \right) + \mathcal{O}(n^{-1})
= V_n(\theta_0) - \sum_j^k \frac{\partial V_n(\theta_0)}{\partial \theta_j} e_j \left( M' \tilde{S} M \right)^{-1} M' \tilde{S} G_n(\theta_0) + \mathcal{O}(n^{-1})
\]
\[
W_n^{-1} = V_n(\hat{\theta}) - \sum_j^k \frac{\partial V_n(\theta_0)}{\partial \theta_j} e_j' \left( \tilde{\Upsilon} - \hat{\Upsilon} \right) G_n(\theta_0) + \mathcal{O}(n^{-1})
= V_n(\hat{\theta}) - \sum_j^k \frac{\partial V_n(\theta_0)}{\partial \theta_j} G_n(\theta_0)' \left( \tilde{\Upsilon} - \Upsilon \right)' e_j + \mathcal{O}(n^{-1})
\] (142)

The last line is because \( \hat{\Upsilon} = \Upsilon + \mathcal{O}(n^{-1/2}) \). This is not the case for \( \tilde{\Upsilon} \), because \( \tilde{S} \) is a function of the first round weighting matrix \( \tilde{W} \), which is not necessarily a consistent
estimate of $\Sigma^{-1}$. Substituting equations (141) and (142) into (139) provides the result:

$$S_n(\hat{\theta}) = \Sigma^{-1} - \Sigma^{-1} \left( V_n(\hat{\theta}) - \Sigma \right) \Sigma^{-1}$$

$$+ \Sigma^{-1} \left( (\delta)^2 \frac{1}{n} \sum_i g_i(\hat{\theta})' \Sigma^{-1} G_n(\hat{\theta}) v_i(\hat{\theta}) + (1 - \delta) \sum_{j=1}^k \tilde{\Gamma}_j G_n(\theta_0)'(\tilde{T} - \Upsilon)' e_j \right) \Sigma^{-1}$$

$$+ O(n^{-1})$$

(143)